OIPE JC100
NOV 04 2002
PATENT & TRADEMARK OFFICE

#24 attach Recon

RECEIVED
NOV 07 2002
TECH CENTER 1600/2900

CERTIFICATE OF MAILING
37 C.F.R. § 1.8

I hereby certify that this correspondence is being deposited with the U.S. Postal Service as First Class Mail in an envelope addressed to: Assistant Commissioner for Patents, Washington, D.C. 20231, on the date below:

October 29, 2002
Date

Gina N. Shishima

PATENT

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of:
Susan Lindquist

Serial No.: 09/207,649

Filed: December 8, 1998

For: METHODS FOR IDENTIFYING FACTORS THAT CONTROL THE FOLDING OF AMYLOID PROTEINS OF DIVERSE ORIGIN

Group Art Unit: 1647

Examiner: S. Turner

Atty. Dkt. No.: ARCD:278/GNS

SUPPLEMENTAL APPEAL BRIEF

NOV 0 4 2002
PATENT & TRADEMARK OFFICE

Table of Contents

I. Real Party in Interest 2
II. Related Appeals and Interferences 2
III. Status of the Claims 2
IV. Status of Amendments 3
V. Summary of the Invention 3
VI. Issues on Appeal 4
VII. Grouping of the Claims 4
VIII. Argument 5
A. Rejections Under 35 U.S.C. § 102 Are Improper 5
1. Standard of anticipation 5
2. Hughes *et al.* does not anticipate claims 1, 3, 7-20, and 22 5
3. Claims 8-11, 14-16, 19, 20, and 22 do not stand or fall with the rejected claims 1, 3, 7, 12-13, and 17-18 as being anticipated by Hughes *et al* 10
4. Cordell *et al.* does not anticipate claims 1, 3, 15, 17-19, and 37 12
5. Claims 7-14, 16, 20, 22 and 37 do not stand or fall with the rejected claims 1, 3, 15, 17-19, and 37 as being anticipated by Cordell *et al* 14
6. Findeis *et al.* does not anticipate claims 1, 3, 7, 12-13, 17-18 and 37 16
B. Rejections Under 35 U.S.C. § 112 Are Improper 17
1. Claims 7-11 are enabled under 35 U.S.C. § 112, first paragraph 17
2. Claims 7-11 are definite under 35 U.S.C. § 112, second paragraph 21
C. Objection of Claims 7-11 Is Improper 23
IX. Conclusion 24

APPENDIX A: Pending claims on appeal
APPENDIX B: Exhibits
Exhibit 1: Hughes *et al.*
Exhibit 2: Findeis *et al.*
Exhibit 3: Phizicky and Fields.
Exhibit 4: Cordell *et al.*
Exhibit 5: Patino *et al.*
Exhibit 6: King *et al.*
Exhibit 7: Selvaggini *et al.*
Exhibit 8: Ogawa *et al.*

Exhibit 9: Schilthuis *et al.*
Exhibit 10: Mohler *et al.*

OIPE JC100
NOV 04 2002
PATENT & TRADEMARK OFFICE

RECEIVED
NOV 07 2002
TECH CENTER 1600/2900

PATENT

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of: Susan Lindquist	Group Art Unit: 1647
Serial No.: 09/207,649	Examiner: S. Turner
Filed: December 8, 1998	Atty. Dkt. No.: ARCD:278/GNS
For: METHODS FOR IDENTIFYING FACTORS THAT CONTROL THE FOLDING OF AMYLOID PROTEINS OF DIVERSE ORIGIN	

SUPPLEMENTAL APPEAL BRIEF

Board of Patent Appeals and Interferences
U. S. Patent and Trademark Office
Washington, D.C. 20231

Appellant hereby submits an original and three copies of this Supplemental Appeal Brief to the Board of Patent Appeals and Interferences in response to the Office Action dated April 29, 2002. Appellant is submitting concurrently herewith a request under 37 C.F.R. §1.193(b)(2)(ii) to reinstate the present appeal. The previous Appeal Brief filed on January 24, 2002 is herein incorporated by reference in its entirety.

The fee for filing this Supplemental Appeal Brief of $160.00 was previously paid prior to the examiner's re-opening of prosecution. The credit for the initial Appeal Brief should apply hereto.

A request for a three-month extension of time is also included herewith, along with the required fee. This extension of time will bring the due date to October 29, 2002. Should any other fees be due, or the attached fee be deficient or absent, the Commissioner is authorized to withdraw the appropriate fees from Fulbright & Jaworski Deposit Account No. 50-

1212/ARCD:278. Please date stamp and return the enclosed postcard to evidence receipt of this document.

I. REAL PARTY IN INTEREST

The real party in interest is the assignee, ARCH Development Corporation.

II. RELATED APPEALS AND INTERFERENCES

There are no interferences or appeals for related cases.

III. STATUS OF THE CLAIMS

Claims 1-36 were originally filed in the application on December 8, 1998. Claims 1-22, the Group I claims, were elected for prosecution in telephonic conference with the Examiner on July 7, 1998. Claims 1, 3, 4, 7, 12, 15, and 16 were amended, claim 37 was added, and claims 2, 5, 6, and 21 were cancelled without prejudice or disclaimer in the Response to Office Action dated August 4, 1999. Claims 22-36 were cancelled without prejudice or disclaimer, claims 1 and 4 were amended, and claims 38-40 were added in the Response to Office Action dated February 1, 2000. Claim 4 was cancelled without prejudice or disclaimer and claims 1, 7, 12, and 14 were amended in Response to Office Action dated October 24, 2000. Claims 1, 3, 7-20, 22, and 37-40 were rejected in the Final Office Action dated June 19, 2001. Claims 38-40 were withdrawn from consideration as being directed to a non-elected invention in the Final Office Action dated June 19, 2001. Claims 1, 3, 7-20, 22, and 37 were rejected in the Office Action dated April 29, 2002 (the "Action").

Accordingly, claims 1, 3, 7-20, 22, and 37 are pending. Of these, claims 1, 3, 7-20, 22, and 37 are the subject of the present appeal and stand appealed. A copy of the appealed claims is attached as Appendix A to this Brief.

IV. STATUS OF AMENDMENTS

No amendments have been filed subsequent to the final rejection.

V. SUMMARY OF THE INVENTION

The present invention discloses a method of identifying a candidate substance that inhibits the aggregation of a mammalian aggregate-prone amyloid protein, comprising: (a) contacting a yeast cell that expresses a chimeric aggregate-prone amyloid protein comprising a mammalian aggregate-prone amyloid peptide with said candidate substance under conditions effective to allow aggregated amyloid formation; and (b) determining the ability of said candidate substance to inhibit the aggregation of the aggregate-prone amyloid protein. In other embodiments, the mammalian aggregate-prone amyloid protein comprises a PrP or β-amyloid polypeptide. In further embodiments the chimeric protein comprises at least an aggregate forming domain of a mammalian aggregate-prone amyloid protein operably attached to a detectable marker protein. In certain embodiments, the marker protein is green fluorescent protein, luciferase, a drug-resistance marker protein, or a hormone receptor. In particular embodiments, the hormone receptor is a glucocorticoid receptor. In other embodiments, the chimeric protein comprises at least an aggregate forming domain of PrP or β-amyloid. In certain embodiments, the chimeric protein comprises at least about amino acids 1-42 of β-amyloid protein. In other embodiments, the chimeric protein comprises Sup35 in which the N-terminal domain has been replaced by amino acids 1-42 of β-amyloid protein. In further embodiments, any aggregation of the mammalian aggregate-prone amyloid protein is detected by the ability of the aggregated protein to bind Congo Red or is detected by increased protease resistance of the aggregated protein. In other embodiments, the aggregate-prone amyloid protein is labeled. The label may be a radioactive isotope, a fluorophore, or a chromophore. In certain embodiments,

the label is ^{35}S. In particular embodiments, the fluorophore comprises a green fluorescent protein polypeptide. In other embodiments, the yeast cell overexpresses Hsp104. In further embodiments, the aggregated amyloid formation is evidenced by the formation of fibrillary material. Specification at page 5, lines 10-15, 18-19, and 21-30; page 6, lines 1-30; page 7, lines 1 and 8-9.

VI. ISSUES ON APPEAL

A. Whether claims 1, 3, 7-20, and 22 are properly rejected under 35 U.S.C. § 102(b) as being anticipated by Hughes *et al.* (Exhibit 1)?

B. Whether claims 1, 3, 7-20, 22, and 37 are properly rejected under 35 U.S.C. §102(b) as being anticipated by Cordell *et al.* (Exhibit 4)?

C. Whether claims 1, 2, 7, 12-13, 17-18, and 37 are properly rejected under 35 U.S.C. § 102(e) as being anticipated by Findeis *et al.* (Exhibit 2)?

D. Whether claims 7-11 are properly rejected under 35 U.S.C. § 112, first paragraph as being not enabled?

E. Whether claims 7-11 are properly rejected under 35 U.S.C. § 112, second paragraph as being indefinite for failing to particularly point out and distinctly claim the subject matter which Appellant regards as the invention?

F. Whether claims 7-11 are properly objected to under 37 C.F.R. § 1.75(c) as being of improper dependent form for failing to further limit the subject matter of a previous claim?

G. Whether claims 1, 3, 7-20, and 37 are properly rejected under 35 U.S.C. § 112, second paragraph, as indefinite based on the terms "aggregate-prone amyloid protein/peptide," "aggregated amyloid formation," "aggregation," "aggregate-forming domain," "β-amyloid," "β-amyloid," "β-amyloid protein," and "chimeric" as being repugnant in the art?

H. Whether claims 1, 7, 8, 17-18, 20, and 22 are properly rejected under 35 U.S.C. § 102(b) as being anticipated by Patino *et al.* (Exhibit 5)?

I. Whether claim 16 is properly rejected under 35 U.S.C. § 103(a) as being obvious over Hughes *et al.*, Cordell *et al.*, Findeis *et al.*, and Patino *et al.*, in view of King *et al.* (Exhibit 6) and Selvaggini *et al.* (Exhibit 7)?

J. Whether claims 9-11 are properly rejected under 35 U.S.C. § 103(a) as being obvious over Hughes *et al.*, Findeis *et al.*, and Patino *et al.*, in view of Ogawa *et al.* (Exhibit 8), Schilthuis *et al.* (Exhibit 9), and Mohler *et al.* (Exhibit 10)?

VII. GROUPING OF THE CLAIMS

For purposes of this Appeal, the claims do not stand or fall together as set forth in the Argument below.

Claims 15, 19 and 37 stand separately from the other claims because they have not been rejected as anticipated by the Hughes *et al.* reference, and thus, even if the Hughes *et al.* reference is found to be anticipating of some of the claims, claims 15, 19, and 37 are potentially patentable over it.

Claim 19 is also separately patentable from the other claims because it has not been rejected as anticipated by the Hughes *et al.* reference, the Findeis *et al.* reference, or the Patino reference. Thus, even if any of these references if found to be anticipating of some of the claims, claim 19 is potentially patentable over them.

Claim 7 stands separately from the other claims because it has not been rejected as anticipated by the Cordell *et al.* reference, and thus, even if the Cordell *et al.* reference is found anticipating of some of the claims, claim 7 is potentially patentable over it.

Claim 3 stands separately from the other claims because it has not been rejected as anticipated by the Patino *et al.* reference, and thus, even if the Patino *et al.* reference is found anticipating of some of the claims, claim 3 is potentially patentable over it.

Claims 20 and 22 stand separately from the other claims because they have not been rejected as anticipated by the Hughes *et al.* reference, the Findeis *et al.* reference, or the Cordell reference, and thus, even if these references are found to anticipate some of the claims, claims 20 and 22 are potentially patentable over them.

Claim 16 stands separately from the other claims because it has not been rejected as anticipated. Even if a cited reference is found anticipating of some of the claims, claim 16 is potentially patentable over those references.

Claims 9-11 stand separately from the other claims because they have not been rejected as anticipated. Even if a cited reference is found anticipating of some of the claims, claims 9-11 are potentially patentable over those references. Furthermore, the references cited against these claims are different than the references cited against claim 16 under an obviousness rejection.

In any event, every element of these rejected claims are not disclosed either expressly or inherently in the cited references.

VIII. ARGUMENT

A. Substantial evidence required to uphold the examiner's position

As an initial matter, Appellant notes that findings of fact and conclusions of law by the U.S. Patent and Trademark Office must be made in accordance with the Administrative Procedure Act, 5 U.S.C. § 706(A), (E), 1994. *Dickinson v. Zurko*, 527 U.S. 150, 158 (1999). Moreover, the Federal Circuit has held that findings of fact by the Board of Patent Appeals and Interferences must be supported by "substantial evidence" within the record. *In re Gartside*, 203 F.3d 1305, 1315 (Fed. Cir. 2000). In *In re Gartside*, the Federal Circuit stated that "the 'substantial evidence' standard asks whether a reasonable fact finder could have arrived at the agency's decision." *Id.* at 1312.

Accordingly, it necessarily follows that an Examiner's position on Appeal must be supported by "substantial evidence" within the record in order to be upheld by the Board of Patent Appeals and Interferences.

B. Rejections Under 35 U.S.C. § 102 Are Improper

1. Standard of anticipation

The novelty of the claimed invention is tested by determining whether or not the claimed invention is anticipated by the prior art as defined in 35 U.S.C. § 102. Anticipation requires that each and every element of the claimed invention be described, either expressly or inherently, in a single prior art reference. *Telemac Cellular Corp. v. Topp Telecom, Inc.*, 247 F.3d 1316, 1327, 58 U.S.P.Q.2d 1545, 1552 (Fed. Cir. 2001); *Verdegaal Bros., Inc. v. Union Oil Co.*, 814 F.2d 628, 631, 2 U.S.P.Q.2d 1051, 1053 (Fed. Cir. 1987). An anticipation analysis requires identifying the elements in the claims, determining their meaning in light of the specification and prosecution history, and identifying the corresponding elements disclosed in the anticipating reference. *Helifix Limited v. Blok-Lok, Ltd.*, 208 F.3d 1339, 1346, 54 U.S.P.Q.2d 1299, 1303 (Fed. Cir. 2000); *Lindemann Maschinenfabrik v. American Hoist & Derrick Co.*, 730 F.2d 1452, 1458, 221 U.S.P.Q. 481, 485 (Fed. Cir. 1984).

2. Hughes *et al.* does not anticipate claims 1, 3, 7, 12-13 and 17-18

The Office Action dated April 29, 2002 rejects claims 1, 3, 7, 12-13, and 17-18 under 35 U.S.C. § 102(b) as being anticipated by Hughes *et al.*, PNAS, 93:2065-70, 1996 ("Hughes *et al.*"). Specifically, the Action contends that Hughes *et al.* clearly measures the interaction of Aβ peptides (β-amyloid peptides) that aggregate, as evidenced in Figure 1 of this reference. The Action reasons that the Appellant has not defined "aggregates" to exclude the interaction of Aβ monomers and that the direct interaction of these monomers constitutes an "aggregation" as depicted in Figure 1. The Action states that the conditions in Hughes *et al.* appear to be sufficient for aggregation and that the Aβ TT mutant in this reference is an identified substance which inhibits aggregation.

The Action further argues that claims 3-4 are anticipated by this reference because the aggregate-prone protein comprises β amyloid chimeric proteins. Claim 7 is alleged to be anticipated as the chimeric protein comprises an aggregate forming domain operably attached to a detectable marker protein. Claims 12-13 are alleged to be anticipating because the amyloid polypeptide is β amyloid residues 1-42. Claims 17-18 are alleged to be anticipated as the aggregate is labeled with a chromophore. From this, the Action concludes that Hughes *et al.* anticipates the claimed invention.

Appellant respectfully traverses this rejection. Hughes *et al.* does not anticipate claims 1, 3, 7-20, and 22 of Appellant's presently claimed invention.

As stated above, anticipation requires that each and every element of the claimed invention be described, either expressly or inherently, in a single prior art reference. *Telemac Cellular Corp. v. Topp Telecom, Inc.*, 247 F.3d 1316, 1327, 58 U.S.P.Q.2d 1545, 1552 (Fed. Cir. 2001).

Appellant claims as the present invention "A method of identifying a candidate substance that inhibits the aggregation of a mammalian aggregate-prone amyloid protein, comprising: (a) contacting a yeast cell that expresses a chimeric aggregate-prone amyloid protein comprising a mammalian aggregate-prone amyloid peptide with said candidate substance under conditions effective to allow aggregated amyloid formation; and (b) determining the ability of said candidate substance to inhibit the aggregation of the aggregate-prone amyloid protein. Claim 1.

The Hughes *et al.* reference (Exhibit 1) makes it clear that the authors are not investigating aggregation, but rather the interaction of monomers. In the opening paragraphs of the article, the rationale for the studies in the article are described and stated as:

> Kinetic studies on AB aggregation have demonstrated that amyloid formation is a nucleation-dependent phenomenon, and that lag time precedes aggregation, the length of which may depend on protein concentration. The nucleation even may

> therefore be the rate-determining step of *in vivo* amyloidosis...Interaction between two monomers, a thermadynamically unfavorable intermolecular interaction, may be a critical step in nucleation.

Hughes *et al.*, page 2065, column 2 (citations omitted). It is evident that monomer formation is hypothesized as part of a nucleation event that ***precedes aggregation***. Hughes *et al.* goes on to state that "[t]he slow and thermodynamically unfavorable interactions between individual monomers may be the rate-limiting step in aggregation" (page 2070, column 1). A person skilled in the art could only understand this to mean that the interaction of two monomers is necessary ***prior to the aggregation*** of multiple monomers. As evidenced by Hughes *et al.*, a person skilled in the art would not interpret the term "aggregation" to include the interaction between two individual monomers. Furthermore, Appellant's specification does not define the term "aggregation" to include the interaction between two individual monomers. As stated above, an anticipation analysis requires identifying the elements in the claims, determining their meaning in light of the specification and prosecution history, and identifying the corresponding elements disclosed in the anticipating reference. *Helifix Limited v. Blok-Lok, Ltd.*, 208 F.3d 1339, 1346, 54 U.S.P.Q.2d 1299, 1303 (Fed. Cir. 2000). Since a person skilled in the art would not interpret the presently pending claims as encompassing monomer interactions when read in light of Appellant's specification, the Hughes *et al.* reference does not anticipate the present invention.

Furthermore, the examiner has improperly assumed that an aggregate includes the interaction of two monomers. There is no reason why this assumption is valid. Because the Appellant's specification does not support this definition, the burden rests on the Examiner to provide evidence that the interaction of two monomers constitutes "aggregation." The Examiner has not met this burden. As stated directly above, the very reference the Examiner cites to support this proposition, Hughes *et al.*, specifically distinguishes between the interaction of two

monomers and AB ***aggregation*** (page 2065, column 2). Furthermore, in contending that the interaction of AB monomers in Hughes *et al.* constitutes "aggregation," the Examiner refers to Figure 1. Appellant respectfully points out that the authors of Hughes *et al.* admit that Figure 1 is merely a "schematic representation." Hughes *et al.*, at 2066, legend to Figure 1. Furthermore, the occurrence of ***all*** the events in this figure are not supported by any experimental data and does not constitute evidence that the reference meets the "under conditions effective to allow aggregated amyloid formation." Thus, the Examiner has not put forth "substantial evidence" to support the rejection of claims 1, 3, 7, 12-13 and 17-18 as being anticipated by Hughes *et al.*

Moreover, another reference cited by the Examiner, Findeis *et al.*, (Exhibit 2) also uses the term "aggregation" to describe the interaction of multiple peptides and ***not between just two peptides***. For example, Findeis *et al.* states that "[t]he term 'aggregation of B amyloid peptides' refers to a process whereby the peptides associate with each other to form a ***multimeric***, largely insoluble complex." Column 8, lines 11-16. These citations provide evidence that a person skilled in the art is well aware that the formation of a multimeric complex includes more than just the interaction between two peptides. Thus, there is no reason why a person of skill in the art would understand that the interaction of two monomers constitutes "aggregation." Since a person skilled in the art does not understand the term "aggregation" to describe the interaction of two monomers and since the Appellant has not defined "aggregation" to include such a description, Hughes *et al.*, which is said to discuss monomer interactions, does not anticipate the presently claimed invention.

Furthermore, Hughes *et al.* describes a typical use of the yeast two-hybrid system. In fact, Figure 1 of this reference precisely shows that the system can evaluate the ability of ***only*** monomers to associate. The steps of Appellant's claimed invention specifically evaluate "aggregated amyloid formation." In fact, the assay in Appellant's claimed (Claim 1) invention is

set up "under conditions effective to allow aggregated amyloid formation." In contrast, the two-hybrid system in Hughes *et al.* relies on the ability of two molecules to interact in the nucleus to promote transcription of a reporter gene, which is subsequently assayed. The Hughes *et al.* reference states, "The yeast system described in this paper offers an opportunity to study the interaction of monomeric AB peptides....This system may therefore provide an opportunity to freeze-frame the monomer-monomer interaction." Hughes *et al.* p. 2070, column 1. Figure 1 in Hughes *et al.* also confirms this by stating: "The system therefore provides an opportunity to examine interaction between two monomeric AB molecules, an essential first step in the nucleation event leading to fibril formation." Hughes *et al.* page 2066. Appellant's specification states that "amyloid or amyloid like deposits are generally insoluble fibrillary material." Specification at page 5, lines 18-19. This indicates not only that the method of Hughes *et al.* is not performed under conditions to promote "aggregated amyloid formation," which is a limitation recited by the claims, but also that it ***would not be*** performed under such conditions because such aggregation of the monomers would prevent transcription—the very event being assayed—from occurring at all.

Also, nowhere in Hughes *et al.* is there a suggestion that the assay be performed under conditions to promote aggregated amyloid formation, which is not surprising given that the assay was not intended to evaluate amyloid formation. The limitation that steps of the assay be performed "under conditions effective to allow aggregated amyloid formation " is conspicuously absent from Hughes *et al.* The yeast two-hybrid system is a widely used, well understood assay that relies on the ability of two polypeptides to bind DNA in a specific (as opposed to non-specific) manner and to associate in a stereo-specific way with proteins that form part of the host cell's transcriptional machinery. This assay is critically dependent on the ability of the polypeptides involved in the assay to reach the nucleus so they can specifically interact with one

another to transcribe the reporter gene being assayed. *See* Phizicky and Fields, *Microbiol. Rev.* 59:94-123, 106 (1995) ("The two-hybrid system is limited to proteins that can be localized to the nucleus, which may prevents its use with certain extracellular proteins.") (Exhibit 3). In the assays of the present invention, aggregation prevents the polypeptides from reaching the nucleus. In the context of the Hughes *et al.* reference, aggregation of a mammalian aggregate-prone amyloid proteins causes them to be insoluble such that the A monomers in the two-hybrid system would be unable to associate in the nucleus, and they would be unavailable to promote transcription. Therefore, the yeast-two hybrid system is simply inoperable for the intended purpose of the assay if practiced according to the limitations of the claimed invention. Thus, Hughes *et al.* does not suggest that the assay be performed under conditions to promote aggregated amyloid formation.

Since the Hughes *et al.* reference does not teach or suggest all of the claim limitations in Appellant's claimed invention, it does not anticipate the presently claimed invention.

Accordingly, Appellant respectfully requests that the rejection of claims 1, 3, 7-20, and 22 under 35 U.S.C. § 102(b) as being anticipated by Hughes *et al.* be withdrawn.

3. Claims 8-11, 14-16, 19, 20, and 22 do not stand or fall with the rejected claims 1, 3, 7, 12-13, and 17-18 as being anticipated by Hughes *et al.*

The Action rejects claims 1, 3, 7, 12-13, and 17-18 under 35 U.S.C. § 102(b) as being anticipated by Hughes *et al.* The previous Action rejected claims 8-11, 14-16, 19, 20, and 22 on the grounds that these claims depend from the rejected base claims. This rejection appears to have been withdrawn in the present Action. To the extent it has not been withdrawn, Appellant refers to the argument they previously filed in its initial Appeal Brief.

4. Cordell *et al.* does not anticipate claims 1, 3, 15, 17-19, and 37

The Action rejects claims 1, 3, 12-13, 15, 17-19, and 37 under 35 U.S.C. § 102(b) as being anticipated by Cordell *et al.*, WO 91/04339 ("Cordell *et al.*"). The Action contends that Cordell *et al.* discloses assays and reagents for amyloid deposition including the identification of agents that inhibit amyloid formation. The Action further states that Appellant's definition of "chimeric" includes modified beta amyloid proteins with one or more substituted amino acids, and thus, the substituted mutants of Cordell *et al.* are chimeric aggregate-prone amyloid proteins. This references is also said to teach that the amyloid products, including β-amyloid 1-42, may be expressed in yeast, that the compounds are detected by Congo red, thioflavin S, or silver salt staining, and that labels may be employed.

Appellant respectfully traverses this rejection. Cordell *et al.* does not anticipate claims 1, 3, 12-13, 15, 17-19, and 37 of Appellant's presently claimed invention.

Appellant takes the position that Cordell *et al.* (Exhibit 4) does not teach or suggest each and every element of the claimed invention. In contrast to what Appellant claims, Cordell *et al.* discloses a single amyloid polypeptide with substitutions of one or more amino acids. Cordell *et al.* does ***not*** teach a "chimeric aggregate prone amyloid protein" anywhere in the disclosure which is a required element of claim 1 in the present invention. Furthermore, there is no suggestion to employ the use of a chimeric protein in Cordell *et al.* The Action's contention that the substitution of one or more amino acids in a single amyloid peptide constitutes a chimeric protein is incorrect. It is well known to people skilled in the art that a chimeric protein contain at least two separate polypeptides combined into one whole protein. In fact, Appellant's specification defines "chimeric protein" to mean "the protein comprises *polypeptides* that do not naturally occur together in a single protein unit." Specification at page 5, lines 26-27. Appellant's specification uses the term "polypeptides" and it is also used in the plural form. The

definition does not say a chimeric has a substitution of a single amino acid, as the Action contends. This shows an explicit intent to define a "chimeric protein" as including ***at least two*** separate polypeptides that do not naturally occur together. In stark contrast, the Action contends that a ***single*** polypeptide with a substituted amino acid is considered a "chimeric protein." When Appellant's claims are read in light of the specification, a "chimeric protein" comprises of ***at least two*** separate polypeptides that do not naturally occur together. Since Cordell *et al.* does not teach or suggest a "chimeric aggregate-prone amyloid protein" as defined by the specification, it does not teach every element of the presently claimed invention. Therefore, Cordell *et al.* does not anticipate claims 1, 3, 7-20, 22, and 37.

Furthermore, some of the dependent claims are also not anticipated by Cordell *et al.* for other reasons. Claim 12 states "The method of claim 1, wherein the chimeric protein comprises at least an aggregate forming domain of PrP or β-amyloid." Claim 13 depends from claim 12. Thus the limitation in claim 12 extends to claim 13. Cordell *et al.* does not teach or suggest "...an aggregate forming domain of PrP or β-amyloid."

Claim 16 states "The method of claim 1, wherein any aggregation of the mammalian aggregate-prone amyloid protein is detected by increased protease resistance of the aggregated protein." Claim 20 states "The method of claim 18, wherein the flourophore comprises a green fluorescent protein polypeptide" and claim 22 claims "The method of claim 1, wherein said yeast cell over expresses Hsp 104." Finally, claim 37 claims "The method of claim 1, wherein aggregated amyloid formation is evidenced by the formation of fibrillary material." Cordell *et al.* does not teach or suggest any of the limitations claimed in claims 16, 20, 22, or 37. In fact, the Action concedes this by stating that the sole basis for rejecting these claims is that they depend from rejected base claims.

Accordingly, Appellant respectfully requests that the rejection of claims 1, 3, 7-20, 22, and 37 under 35 U.S.C. § 102(b) as being anticipated by Cordell *et al.* be withdrawn.

5. Claims 7-14, 16, 20, 22 and 37 do not stand or fall with the rejected claims 1, 3, 15, 17-19, and 37 as being anticipated by Cordell *et al.*

The Action rejects claims 1, 3, 7-20, 22, and 37 under 35 U.S.C. § 102(b) as being anticipated by Cordell *et al.* The previous Final Office Action rejected claims 7-14, 16, 20, 22, and 37 on the grounds that these claims depend from rejected base claims. This rejection appears to have been withdrawn in the present Action. To the extent it has not been withdrawn, Appellant refers to the argument they previously filed in its initial Appeal Brief.

6. Findeis *et al.* does not anticipate claims 1, 3, 7, 12-13, 17-18 and 37

The Action rejects claims 1, 3, 7, 12-13, 17-18, and 37 under 35 U.S.C. § 102(e) as being anticipated by U.S. Patent No. 5,854,204 to Findeis *et al.* The Action contends that Findeis *et al.* teaches A-beta peptides including chimeric peptides, which differ from naturally occurring beta amyloid sequences and screening assays to identify modulatory influences on amyloid aggregation.

Appellant respectfully traverses this rejection. Findeis *et al.* does not anticipate claims 1, 3, 7, 12-13, 17-18, and 37.

As noted above, Appellant claims "A method of identifying a candidate substance that inhibits the aggregation of a mammalian aggregate-prone amyloid protein, comprising: (a) contacting a yeast cell that expresses a chimeric aggregate-prone amyloid protein comprising a mammalian aggregate-prone amyloid peptide with said candidate substance under conditions effective to allow aggregated amyloid formation; and (b) determining the ability of said candidate substance to inhibit the aggregation of the aggregate-prone amyloid protein." Claim 1.

In contrast, Findeis *et al.* (Exhibit 2) never teaches or suggests performing the screening assay in yeast. This is evident in light of Example 5 in the Findeis *et al.* reference (column 51, lines 25 to column 53, line 27). Example 5 in Findeis *et al.* employs the use of a seeded static assay and a shaken plate aggregation assay to identify B-Amyloid modulators. Findeis *et al.* discloses that these assays simply mix a candidate modulator with AB-monomers to determine the affects of the candidate modulator. In fact, these screening assays employed by the Findeis *et al.* reference do not employ the use of yeast cells. In stark contrast, Appellant's screening assay comprises "***contacting a yeast cell*** that expresses a chimeric aggregate-prone amyloid protein comprising a mammalian aggregate-prone amyloid peptide with said candidate substance under conditions effective to allow aggregated amyloid formation." Thus, Findeis *et al.* does not teach or suggest every element claimed in Appellant's presently claimed invention.

Furthermore, Findeis *et al.* fails to teach or suggest ***incubating*** yeast cells "under conditions effective to allow aggregated amyloid formation," an element expressly required in Appellant's claimed invention. As stated directly above, the screening assays used to identify β-amyloid modulators do not employ the use of yeast cells. Thus, Findeis *et al.* fails to teach or suggest every element required in Appellant's claimed invention either expressly or inherently. Thus, Findeis *et al.* does not anticipate claims 1, 3, 7, 12-13, 17-18, and 37.

Moreover, Findeis *et al.* does not render claims 1, 3, 7, 12-13, 17-18, and 37 obvious. This is evident in light of the disclosure in Example 5 in Findeis *et al.* Example 5 of Findeis *et al.* teaches away from using a yeast cell to perform screening assay for inhibitors of B-Amyloid aggregation. As discussed above, Example 5 in Findeis *et al.* employs the use of screening assays that simply mix a candidate modulator with AB-monomers to determine the effects of the candidate modulator. Also stated above, these assays employed by Findeis *et al.* do not employ

the use of yeast cells. Since the assays employed by Findeis *et al.* teach away from employing the use of yeast cells, Findeis *et al.* does not render claims 1, 3, 7, 12-13, 17-18, and 37 obvious.

Accordingly, Appellant respectfully requests that the rejection of claims 1, 3, 7, 12-13, 17-18, and 37 under 35 U.S.C. § 102(e) as being anticipated by Findeis *et al.* be withdrawn.

7. Patino *et al.* does not anticipate claims 1, 7, 8, 17-18, 20, and 22

The Action rejects claims 1, 7, 8, 17-18, 20, and 22 under 35 U.S.C. § 102(b) as anticipated by Patino *et al.* (Exhibit 5). The Action contends that Patino *et al* teaches analysis of the aggregation and inhibition of yeast cells expressing Sup35-green fluorescent protein labeled chimeric peptides in yeast cells and in yeast cells overexpressing Hsp104. It further contends that Hsp104 is a candidate agent that inhibits aggregation. Finally, it argues that Sup35 forms amloid-like aggregates and aggregates in yeast cells to anticipate the claimed invention. Appellant respectfully traverses this rejection.

Appellant notes that this rejection over Patino *et al.* is the ***same*** rejection that was previously withdrawn by the examiner in the Office Action dated June 19, 2001. No reason was provided in the most recent Action indicating the reappearance of this rejection.

Patino *et al.* does not disclose a "chimeric aggregate-prone protein comprising a *mammalian* aggregate-prone amyloid peptide." Patino *et al.* concerns Sup35, which is a yeast polypeptide. It does not disclose all of the elements of the claimed invention.

In previous office actions the examiner has suggested there is sequence similarity between yeast prions and mammalian sequences. There is no sound basis for asserting that references that describe yeast prions necessarily disclose protein sequences from *mammalian* aggregate-prone proteins. If the grounds for this rejection are to be maintained, the Examiner is requested to provide evidence in the form of a reference or affidavit that supports the contention

that yeast prion sequences are so similar to mammalian aggregate prone proteins that they are virtually identical. *See* MPEP § 2144.03; 37 C.F.R. § 1.104(d)(2)..

In fact, the specification teaches that PrP, a mammalian prion protein, and Sup 35, a yeast protein that forms aggregates, have **no** sequence identity and are functionally unrelated. Specification page 34, lines 6-8. Sup35 is the subject matter of the Patino *et al.* paper. Thus, the statement in the specification calls into question the Action's assertion that the sequences of the cited art and the subject matter of the invention are similar.

Patino *et al.* does not disclose recited elements of the claimed invention. Accordingly, it does not anticipate the claimed invention. Appellant respectfully requests this rejection be withdrawn.

C. Rejections Under 35 U.S.C. § 103 Are Improper

1. Standard of Obviousness

In order to establish a *prima facie* case of obviousness, three basic criteria must be met: (1) there must be some suggestion or motivation, either in the references themselves or in the knowledge generally available to one of ordinary skill in the art, to modify the reference or to combine reference teachings; (2) there must be a reasonable expectation of success; and (3) the prior art reference (or references when combined) must teach or suggest all the claim limitations. *Manual of Patent Examining Procedure* § 2142. Moreover, the teaching or suggestion to make the claimed combination and the reasonable expectation of success must both be found in the prior art, and not based on Applicant' s disclosure. *In re Vaeck*, 947 F.2d 488, 20 U.S.P.Q. 2d 1438 (Fed Cir. 1991). When "the motivation to combine the teachings of the references is not immediately apparent, it is the duty of the examiner to explain why the combination of the teachings is proper." MPEP § 2142. Without any of these criterion, the burden does not shift to the applicant.

Moreover, "[i]t is impermissible within the framework of 35 U.S.C. § 103 to pick and choose from any on reference only so much of it as will support a given position to the exclusion of other parts necessary to the full appreciation of what such reference fairly suggests to one skilled in the art." *Bausch & Lomb, Inc. v. Barnes-Hind/Hydrocurve, Inc.*, 230 U.S.P.Q. 416 (Fed. Cir. 1986). Moreover, the Federal Circuit has said, "The mere fact that references can be combined or modified does not render the resultant combination obvious unless the prior art also suggests the desirability of the combination." MPEP § 2143.01 citing *In re Mills,* 916 F.2d 680, 16 USPQ2d 1430 (Fed. Cir. 1990).

2. No joint inventors

The Action contends that the application names joint inventors. The sole inventor in this case is Dr. Susan Lindquist. No other inventors have been identified. Therefore, the paragraph regarding an obligation under 37 C.F.R 1.56 to point out the inventor and invention dates of each claim not commonly owned is completely inapplicable to the present case.

3. Claim 16 Is Not Rendered Obvious

The Action rejects claim 16 under 35 U.S.C. § 103(a) as being unpatentable over Hughes *et al.*, Cordell *et al.*, Findeis *et al.*, and Patino *et al.* in view of King *et al.* (Exhibit 6), and Selvaggini *et al.* (Exhibit 7). The Action contends that the references of Hughes *et al.*, Cordell *et al.*, Findeis *et al.*, and Patino *et al.* teach the method of claim 1. It further alleges that the references of King *et al.* and Selvaggini *et al.* each teach analysis of aggregation of aggregate prone amyloid proteins via protease resistance. The Action contends that one of skill in the art would have been motivated to make such modification based on the ease of the assay and expectation of positive results using the assay as taught by King and Selvaggini.

Appellant respectfully traverses this rejection. As stated above, the references of Hughes *et al.*, Cordell *et al.*, Findeis *et al.*, and Patino *et al.* do not anticipate the method of claim 1. These references do not teach every element of claim 1. They do not teach "A method of identifying a candidate substance that inhibits the aggregation of a mammalian aggregate-prone amyloid protein, comprising: (a) contacting a yeast cell that expresses a chimeric aggregate-prone amyloid protein comprising a mammalian aggregate-prone amyloid peptide with said candidate substance under conditions effective to allow aggregated amyloid formation; and (b) determining the ability of said candidate substance to inhibit the aggregation of the aggregate-prone amyloid protein." Claim 1. More specifically, they do not teach a screening method involving "contacting a yeast cell that expresses a chimeric aggregate-prone amyloid protein comprising a mammalian aggregate-prone amyloid peptide with said candidate substance under conditions effective to allow aggregated amyloid formation." Moreover, the other cited references of King *et al.* and Selvaggini *et al.* do not fill in the gaps in the disclosures. Accordingly, a proper *prima facie* case of obviousness has not been made because the combination of references does not teach every element of the claimed invention.

Furthermore, there is no motivation to combine references. Appellant objects to the rejection because the Action throws out four primary references and says they can be combined with two other references to render the claimed invention obvious. The Action does not explain how each combination of references supports the grounds for the rejection. Appellant notes that the rejection constitutes approximately eight different rejections, assuming the rejection was based on one primary reference with one secondary reference.

Moreover, there is no motivation to combine references because the contention that a person of ordinary skill in the art would be motivated to use the assay of King *et al.* or Selvaggini *et al.* with any of the primary references is unsupported. First, the primary references

do not disclose screening assays for inhibitors of aggregation, and thus, there is no motivation to use the protease assays described in the secondary references. Second, none of the secondary references mention yeast, and thus, there is no motivation to use the yeast screens of the invention with the allegedly disclosed protease resistance assays in the secondary references. Finally, there is no reason a person of ordinary skill in the art would simply take those portions of the cited references to arrive at what the Action contends is the claimed invention. This amounts to an arbitrary picking and choosing the relies on the instant specification to arrive at the invention. Therefore, there is no motivation or suggestion to combine these references.

Claim 16 is not rendered obvious by the cited references. Appellant respectfully requests this rejection be withdrawn.

4. Claim 9-11 Are Not Rendered Obvious

The Action rejects claims 9-11 under 35 U.S.C. § 103(a) as being unpatentable over Hughes *et al.*, Findeis *et al.*, and Patino *et al.* in view of Ogawa *et al.* (Exhibit 8), Schilthuis *et al.* (Exhibit 9), and Mohler *et al.* (Exhibit 10). The Action contends that the references of Hughes *et al.*, Findeis *et al.*, and Patino *et al.* teach the method of claims 1 and 7. It admits that none of these references teaches a marker protein that is a drug resistance marker protein, a hormone receptor protein, or a glucocorticoid receptor protein. However, the Action further alleges that the references of Ogawa *et al.* teaches a chimeric fusion protein marker involving a green fluorescent protein linked to glucocorticoid receptor. It also contends that Schilthuis *et al.* teaches a chimeric fusion protein marker in which detection is achieved using the thyroid hormone receptor for detection via transcription activation. The Action alleges that Mohler *et al.* teaches a chimeric fusion protein marker where detection is achieved using the gene for membrane-bound neomycin phosphotransferase for conferring drug resistance. The Action contends that one of skill in the art would have been motivated to make such modification based

on the detection achieved and expectation of positive results using the detection assays as taught by Ogawa, Shilthuis, and Mohler.

Appellant traverses this rejection. As stated above, the references of Hughes *et al.*, Findeis *et al.*, and Patino *et al.* do not anticipate the methods of claim 1. They also do not anticipate the methods of claim 7, as described above. These references do not teach every element of these claims. They do not teach "A method of identifying a candidate substance that inhibits the aggregation of a mammalian aggregate-prone amyloid protein, comprising: (a) contacting a yeast cell that expresses a chimeric aggregate-prone amyloid protein comprising a mammalian aggregate-prone amyloid peptide with said candidate substance under conditions effective to allow aggregated amyloid formation; and (b) determining the ability of said candidate substance to inhibit the aggregation of the aggregate-prone amyloid protein." Claim 1. More specifically, they do not teach a screening method involving "contacting a yeast cell that expresses a chimeric aggregate-prone amyloid protein comprising a mammalian aggregate-prone amyloid peptide with said candidate substance under conditions effective to allow aggregated amyloid formation." Moreover, the other cited references of Ogawa *et al.*. Schilthuis *et al.*, and Mohler *et al.* do not fill in the gaps in the disclosures. Accordingly, a proper *prima facie* case of obviousness has not been made because the combination of references does not teach every element of the claimed invention.

Furthermore, the Action does not make a proper *prima facie* case of obviousness because there is no motivation to combine the references. There is simply no reason a person of ordinary skill in the art would choose the markers of those particular references in the context of the present invention. Just because the secondary references may have taught the use of these chimeric proteins as marker and achieved positive results does not mean that a person of ordinary skill in the art would have turned to those markers in the context of the present

invention. Again, the present invention concerns screening using yeast cells, and there is nothing about those secondary references that would motivate a person of ordinary skill in the art to use the teachings of any of the secondary references in combination with the primary references. A *prima facie* case has not been made. Accordingly, this rejection should be withdrawn.

D. Rejections Under 35 U.S.C. § 112 Are Improper

1. Claims 7-11 are enabled under 35 U.S.C. § 112, first paragraph

The previous Final Action rejected claims 7-11 under 35 U.S.C. § 112, first paragraph, as not providing enablement for a mammalian aggregate-prone amyloid protein wherein the protein is a chimeric that comprises an aggregate forming domain. This rejection appears to have been withdrawn in the present Action. To the extent it has not been withdrawn, Appellant refers to the argument they previously filed in its initial Appeal Brief.

2. Claims 7-11 are definite under 35 U.S.C. § 112, second paragraph

The Final Action rejected claims 7-11 under 35 U.S.C. § 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter that the Appellant regards as the invention. Specifically, the Action contends that claims 7-11 are indefinite as the skilled artisan is not reasonably apprised of the metes and bounds of "an aggregate forming domain." The Action further states that the skilled artisan has no guidance by which to determine that portion of a chimeric protein which forms "an aggregate forming domain."

This rejection appears to have been withdrawn in the present Action. To the extent it has not been withdrawn, Appellant refers to the argument they previously filed in its initial Appeal Brief.

3. Claims 1, 3, 7-20, 22, and 37 are definite under 35 U.S.C. § 112, second paragraph

The Action rejects claims 1, 3, 7-20, 22, and 37 under 35 U.S.C. § 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter that applicant regards as the invention. It contends that the terms "aggregate-prone amyloid protein" and "aggregate-prone amyloid peptide" are indefinite because they are not the same term and allegedly the artisan cannot discern the difference. It also argues that the terms "β-amyloid," "β-amyloid," and "β-amyloid protein" are indefinite because the recitations are not the same and the artisan could not discern the difference. The Action further contends that the terms "amyloid protein/peptide" "aggregate amyloid formation," "aggregation," and "chimeric" are used to mean "any non-naturally occurring peptide which wither forms amyloid or amyloid-like deposits," which the Action seems to contend is repugnant to the usual meaning of those terms. The Action also contends that the terms "aggregated amyloid formation" and "aggregation" are indefinite because of the definition in the specification.

Furthermore, the Action argues that the term "chimeric" is indefinite and that "chimeric peptides are recognized in the art as fusion proteins." The Action reads the term as encompassing any non-naturally occurring sequences, such as a peptide mutated by deletion, insertion, or substitution, which is not necessarily a fusion protein.

Finally, the Action contends that "aggregate forming domain is indefinite because the specification fails to provide the structural constraints of the domain and fails the clarify the functional requirements of "aggregation." Appellant respectfully traverses this rejection.

As an initial matter, Appellant notes for the record the frustration of having to address this rejection at this point in prosecution. While several indefiniteness rejections have been set forth in previous office actions, a number of these are new. Appellant is frustrated because this case has been in prosecution since 1998, almost 4 years ago, and most of these claim terms have

been present in the claims throughout this prosecution. To receive new indefinite rejections at this juncture serves to impede the expedient prosecution of this case. Nonetheless, Appellant believes they can be overcome.

The standard for definiteness of a claim is whether a person of skill in the art can determine the scope of the invention based on the language of the claims with "a reasonable degree of certainty." *Manual of Patent Examining Procedure* § 2173.02 (citing *In re Wiggins*, 488 F.2d 538, 179 U.S.P.Q. 421 (C.C.P.A. 1973)).

a) "Aggregate-prone amyloid peptide" and "aggregate-prone amyloid peptide"

Claim 1 recites a "yeast cell that expresses a chimeric aggregate-prone amyloid protein comprising a mammalian aggregate prone amyloid peptide." Appellant strongly urges that the terms in the claims are to be understood in the context of the claims and not in a vacuum. It is clear from the claim that the "aggregate-prone amyloid protein" includes the "peptide." As such, there can be no confusion about whether the peptide is larger than the protein. As for their sequence, it is also clear that the "protein" sequence will necessarily include the "peptide" sequence. Therefore, there is no basis for the contention that the claim is indefinite. Appellant specifically requests this rejection be withdrawn.

b) "β-amyloid polypeptide," "β-amyloid," and "β-amyloid protein"

Various claims refer to the same β amyloid protein using synonymous terms. The terms are used consistently to refer to the same compound and there is no basis that a person of ordinary skill in the art would have a different understanding. Furthermore, the usage of those terms in the specification indicates the terms are synonymous.

The standard for definiteness is whether a person of skill in the art can determine the scope of the invention based on the language of the claims with "a reasonable degree of

certainty." The Action provides no evidence to the contrary. Accordingly, this rejection should be withdrawn.

c) The terms are not repugnant

Very confusingly, the Action contends that the terms "amyloid protein/peptide," "aggregated amyloid formation," aggregation," and "chimeric" in the claims are used to mean "any non-naturally occurring peptide which either forms amyloid or amyloid-like deposits." It appears to suggest that this understanding is repugnant to the usual meaning of these terms. It further argues that "the specification fails to delineate that which is 'amyloid-like' and thus the artisan could not readily discern any of the characteristics of amyloid protein interactions amongst those known, are sufficient or required." Appellant traverses this rejection.

The Action seems to indicate that to the extent the definition of "amyloid peptide/protein" refers to a protein capable of forming "an amyloid or amyloid-like deposit," the term covers proteins not recognized by those in the art, making the term "repugnant." Again, there is no support for what a person of ordinary skill in the art would understand the definition to cover; thus, there is no basis for contending that person would find the definition in the specification to be repugnant. None of the terms used is repugnant to its usage in the art. The examiner has not cited a shred of evidence to indicate that the definitions in the specification and claims are "repugnant" to the definitions of those terms in the art.

Furthermore, the specification provides a description of the invention that uses the terms in a manner consistent with their definitions:

> In important embodiments of the present invention, this yeast system is used in methods of identifying a candidate substance that inhibits the aggregation of an aggregate-prone amyloid protein. Such methods comprise contacting a yeast cell that expresses an aggregate-prone amyloid protein with the candidate substance under conditions effective to allow aggregated amyloid formation, and determining the ability of the candidate substance to inhibit the aggregation of the aggregate-prone amyloid protein.

The term "aggregate-prone amyloid protein" is meant to be any protein that is able to form an amyloid or amyloid-like deposit. Amyloid or amyloid like deposits are generally insoluble fibrillary material. Although many proteins are capable of aggregating at high concentrations, aggregate prone amyloid proteins are able to, and often do, aggregate under physiological conditions, such as inside of a cell. Aggregate-prone amyloid proteins include yeast proteins, such as Sup35 and URE3, and mammalian proteins, such as PrP and β-amyloid polypeptide.

Specification at page 5, lines 10-22.

The Action also makes a confusing argument about the art recognizing "various forms of amyloid fibril aggregation" but that the "peptides which are aggregating are not even required to be amyloid peptides." First, the claims are not directed to fibril aggregation. The Examiner has refused to examine claims 38-40, as being drawn to a nonelected invention because they recite "fibril formation." Thus, either these claims should be examined because they are not a separate invention, or the argument about "amyloid fibril agreggation" is irrelevant to the claimed invention. Second, the Action again appears to review the claims in isolated terms and not in the context of the entire claims. The claims recite a "yeast cell that expresses a chimeric aggregate-prone amyloid protein comprising a mammalian aggregate prone amyloid peptide" and "determining the ability of said candidate substance to inhibit the aggregation of the aggregate-prone amyloid protein." The claims already recite "aggregate-prone *amyloid protein"* and thus, the argument that the peptide does not have to be an *amyloid* peptide is erroneous.

d) "Chimeric"

The Action contends that the term "chimeric" is indefinite. It unclearly alleges the term "encompasses any non-naturally occurring sequence" such as a peptide mutated by deletion, insertion, or substitution, which "is not necessarily a fusion protein as recognized by the fusion of two known naturally occurring compounds to make a single non-naturally occurring compound." This understanding is inaccurate.

Again, the examiner offers **no** evidence to support her position. If the examiner is relying on personal knowledge, Appellant requests an affidavit supporting such a position pursuant to 37 C.F.R. § 1.104(d)(2).

Nonetheless, her contention that a chimeric protein necessarily is a fusion protein is not accurate. A fusion protein constitutes a protein in which all or part of another protein has been fused to its carboxy or amino termini. A chimeric protein may encompass such a fusion protein, but it may also encompass a protein in which an interior domain, has been replaced with another protein's interior domain.

Aside from the unsupported and erroneous argument, the specification shows that the definition of "chimeric polypeptide" does not comport with the definition offered in the Action. The specification states: "By 'chimeric protein' it is meant that the protein comprises polypeptides that do not naturally occur together in a single protein unit." Specification at page 5, lines 26-27. As argued previously, the definition says "polypeptides" and does not refer to single amino acids. Furthermore, it provides examples and states that chimeric proteins may include the aggregate forming domain of a mammalian amyloid polypeptide. Specification at page 5, line 30 to page 6, line 1.

There is no reason a person of ordinary skill in the art would read the specification and not be able to determine the scope of the claim with "reasonable certainty" even though it refers to a "chimeric protein." As such, the claims are definite. Appellant respectfully requests this rejection be withdrawn.

e) "Aggregate forming domain"

Claim 7 recites "The method of claim 1, wherein the chimeric protein comprises at least an aggregate forming domain of a mammalian aggregate-prone amyloid protein operably attached to a detectable marker protein." A person of skill in the art could determine that an

"aggregate forming domain" refers to the amino acids of an aggregate-prone amyloid protein that are involved in aggregation. In fact, Appellant's specification states that "In an important embodiment, the chimeric protein comprises Sup35 in which the N-terminal domain has been replaced by amino acids 1-42 of β-amyloid protein." This disclosure eliminates any indefiniteness of what constitutes an "aggregate forming domain."

Furthermore, the fact that one skilled in the art may have to figure out what portions of an aggregate-prone amyloid protein constitute an aggregate forming domain does not render the claims 7-11 indefinite. As mentioned above, the *Manual of Patent Examining Procedure* simply requires that a person of skill in the art be able to determine the scope of the claim. In the present case, claims 7-11 employ a phrase that can be easily understood, "aggregate forming domain." Furthermore, the specification uses that phrase consistent with its use in the claims. Finally, the specification identifies ways of determining whether a protein aggregates. In fact, the specification states that "the aggregation may be detected by its ability to bind Congo Red and show apple green birefringence under polarized light." Specification at page 13, lines 2-3. Thus, not only is the scope of the claims 7-11 are fully ascertainable to a person of ordinary skill in the art and a person of ordinary skill would be able to practice the present invention. Thus, claims 7-11 are not indefinite.

In this case, the claims employ a phrase that can be easily understood, the Specification employs that phrase consistent with its use in the claims, and furthermore, the Specification identifies ways of determining whether a protein aggregates, for example, at page 13, lines 2-3. Thus, not only is the scope of the claim fully ascertainable to a person of ordinary skill in the art, but that person would be able to practice what he understood to be the invention.

Accordingly, Appellant respectfully requests that the rejection of claims 7-11 under 35 U.S.C. § 112, second paragraph be withdrawn.

4. Objection of Claims 7-11 Is Improper

The Action objects to claims 7-11 under 37 C.F.R. 1.75(c) as being of improper dependent form for failing to further limit the subject matter of a previous claim. The Action contends that the recitation "at least an aggregate forming domain" broadens the scope of claim 1, in particular to the mammalian aggregate-prone amyloid protein.

Appellant respectfully traverses this objection. Claim 7 is written in proper dependent form and further limits the subject matter of claim 1.

Claim 1 states "A method of identifying a candidate substance that inhibits the aggregation of a mammalian aggregate-prone amyloid protein, comprising: (a) contacting a yeast cell that expresses a chimeric aggregate-prone amyloid protein comprising a mammalian aggregate-prone amyloid peptide with said candidate substance under conditions effective to allow aggregated amyloid formation; and (b) determining the ability of said candidate substance to inhibit the aggregation of the aggregate-prone amyloid protein." Claim 1. Claim 7 further limits claim 1 by stating "The method of claim 1, wherein the chimeric protein comprises at least an aggregate forming domain of a mammalian aggregate-prone amyloid protein operably attached to a detectable marker protein." Claim 7. Claim 8 further requires that the "marker protein is green flourescent protein or luciferase. Claim 9 further requires that the "marker protein is a drug-resistance marker protein. Claim 10 further requires that the "marker protein is a hormone receptor. Claim 11 further requires that the "hormone receptor is a glucorticoid receptor."

It is evident from the claim language that claim 7 further limits claim 1 in two respects. First, claim 7 further requires "an aggregate forming domain of a mammalian aggregate-prone amyloid protein." Second, claim 7 also requires that the chimeric protein is "operably attached to a detectable marker protein." Both of these limitations are not present in claim 1. It is also

evident that claims 8-10 further limits claim 7 and that claim 11 further limits claim 10. Thus, since claim 7 further limits claim 1, the objection of claim 7 and its dependent claim 8-11 is improper.

Accordingly, Appellant respectfully requests that the objection of claims 7-11 under 37 C.F.R. 1.75(c) be withdrawn.

5. Objection to the Specification Is Improper

The Action objects to the specification for use of the terms identified in the indefinite rejections as being repugnant. While not appealable, Appellant notes that the terms are not repugnant. First and foremost, the Action does not identify any basis for characterizing the definitions in the specification as "repugnant." The word "repugnant" means "arousing disgust or aversion; repulsive" or "contradictory," according to *The American Heritage College Dictionary*, 3d ed. It is improper for the examiner to resort to such extreme language without much basis.

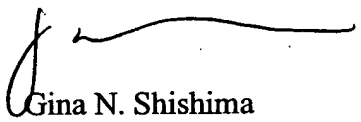
Second, the rejected claims are definite. A person of ordinary skill in the art could determine the scope of the claims with reasonable certainty, with or without the specification because the definitions comport with their ordinary meanings. As such the objections should be withdrawn as well.

II. CONCLUSION

Appellant has provided arguments that overcome the pending rejections. Appellant respectfully submits that the Office Action's conclusions that the claims should be rejected are unwarranted. It is therefore requested that the Board overturn the Action's rejections.

Please date stamp and return the enclosed postcard to evidence receipt of this document.

Respectfully Submitted,



Gina N. Shishima
Reg. No. 45,104
ATTORNEY FOR APPLICANT

FULBRIGHT & JAWORSKI, L.L.P.
600 Congress Ave, Suite 2400
Austin, Texas 78701
(512) 536-3081
(512) 536-4598 (facsimile)

Date: October 29, 2002

OIPE JC100
NOV 04 2002
PATENT & TRADEMARK OFFICE

APPENDIX A
Pending Claims on Appeal

1. A method of identifying a candidate substance that inhibits the aggregation of a mammalian aggregate-prone amyloid protein, comprising:

(a) contacting a yeast cell that expresses a chimeric aggregate-prone amyloid protein comprising a mammalian aggregate-prone amyloid peptide with said candidate substance under conditions effective to allow aggregated amyloid formation; and

(b) determining the ability of said candidate substance to inhibit the aggregation of the aggregate-prone amyloid protein.

3. The method of claim 1, wherein the mammalian aggregate-prone amyloid protein comprises a PrP or β-amyloid polypeptide.

7. The method of claim 1, wherein the chimeric protein comprises at least an aggregate forming domain of a mammalian aggregate-prone amyloid protein operably attached to a detectable marker protein.

8. The method of claim 7, wherein said marker protein is green fluorescent protein or luciferase.

9. The method of claim 7, wherein said marker protein is a drug-resistance marker protein.

10. The method of claim 7, wherein said marker protein is a hormone receptor.

11. The method of claim 10, wherein said hormone receptor is a glucocorticoid receptor.

12. The method of claim 1, wherein the chimeric protein comprises at least an aggregate forming domain of PrP or β-amyloid.

13. The method of claim 12, wherein the chimeric protein comprises as least about amino acids 1-42 of β-amyloid protein.

14. The method of claim 1, wherein the chimeric protein comprises Sup35 in which the N-terminal domain has been replaced by amino acids 1-42 of β-amyloid protein.

15. The method of claim 1, wherein any aggregation of the mammalian aggregate-prone amyloid protein is detected by the ability of the aggregated protein to bind Congo Red.

16. The method of claim 1, wherein any aggregation of the mammalian aggregate-prone amyloid protein is detected by increased protease resistance of the aggregated protein.

17. The method of claim 1, wherein the aggregate-prone amyloid protein is labeled.

18. The method of claim 17, wherein the label is a radioactive isotope, a fluorophore, or a chromophore.

19. The method of claim 18, wherein the label is ^{35}S.

20. The method of claim 18, wherein the fluorophore comprises a green fluorescent protein polypeptide.

22. The method of claim 1, wherein said yeast cell overexpresses Hsp104.

37. The method of claim 1, wherein aggregated amyloid formation is evidenced by the formation of fibrillary material.

BEST AVAILABLE COPY

(see above) with *Eco*RI and *Xho* I. This *Eco*RI–*Xho* I fragment was placed in pJG4-5, a 2-μm *TRP1* plasmid (30, 31), in translational frame with the codons for the simian virus 40 large T nuclear localization signal, the B42 transactivation domain, and the hemagglutinin (HA) epitope tag. The prey fusion protein (16 kDa) will be inducible in yeast grown on minimal medium (MM) containing 2% galactose and 1% raffinose (Gal/Raf) but not in yeast grown on 2% glucose (Glc). Amino acids 3 and 4 (glutamate, phenylalanine) of Aβ (at the point where it is fused to the HA tag) are generated by codons in the *Eco*RI site.

To construct the mutant prey plasmid that contained the AβTT-encoding sequence, a fragment was constructed from a 147-bp oligonucleotide representing the mutation of Phe^{19}-Phe^{20} to Thr^{19}-Thr^{20} within Aβ synthesized on a Millipore model 8909 Expedite nucleic acid synthesis system as follows. The oligonucleotide 5′-AGGCCTGAATTCCGACATGAC-TCAGGATATGAAGTTCATCATCAAAAATTGGT-GACTACTGCAGAAGATGTGGGTTCAAACAAA-GGTGCAATCATTGGACTCATGGTGGGCGGT-GTTGTCATAGCGTAGGTCGACCTCGAGAGGCCT-3′ was annealed with a complementary short oligonucleotide, 5′-AGGCCTCTCGAGGTCGACC-3′, and filled in by Klenow DNA polymerase (BRL). The fragment was extracted with phenol/chloroform and purified with the qiaquick-spin PCR purification kit (Qiagen). The sample was then digested with *Eco*RI and *Xho* I and placed into a ligation reaction mixture in a 7:1 ratio with *Eco*RI/*Xho* I-digested pJG4-5 prey plasmid. The plasmids were propagated and grown in DH5α subcloning-efficiency competent cells from BRL.

The accuracy of the reading frames in the bait and prey plasmids was verified with an automated Applied Biosystems sequencer employing 373 software. Sequences were confirmed to be correct by the analysis features of SEQUENCE EDITOR and MACVECTOR software (data not presented).

Western blot analyses were performed (32) to show that the bait and prey plasmids expressed the expected fusion proteins (data not presented).

Transformation of Strain with Reporter, Bait, and Prey Plasmids. The selection strain was made by transforming the EGY48 yeast strain with a *URA3 lacZ* (β-galactosidase) reporter plasmid and the *HIS3* bait plasmid by the lithium acetate method (27). The yeast selection strain harboring the bait and reporter plasmids was transformed with the prey plasmid DNA (27), and tryptophan utilization phenotype was used (in addition to Ura and His markers for bait and *lacZ* reporter plasmids, respectively) for selection of transformants with prey plasmids.

Determination of Bait–Prey Interaction. Yeast strains containing the appropriate bait and prey plasmids were grown to an OD_{600} of 0.5, diluted 1000-fold, and spotted on plates containing Glc Ura^- His^- Trp^- 5-bromo-4-chloro-3-indolyl β-D-galactopyranoside (X-Gal) medium or Gal/Raf Ura^- His^- Trp^- X-Gal medium to assess the transcriptional activation of the *lacZ* reporter gene. Suitably diluted cell suspensions were also spotted on Gal/Raf Ura^- His^- Trp^- Leu^- medium and Glc Ura^- His^- Trp^- Leu^- medium to assess the transcriptional activation of the *LEU2* gene.

β-Galactosidase Activity in Liquid Cultures of Yeast. Cells were assayed for β-galactosidase activity by the *o*-nitrophenyl β-D-galactopyranoside (ONPG) method (27). The experiment was repeated in triplicate and the plotted data represent an average value of the values for the three samples. The statistical significance was computed with Student's *t* test in a two-tailed analysis.

Immunoprecipitation and Western Blot. Extracts were made from EGY48 cells that contained a prey plasmid encoding B42-Aβ and a bait plasmid encoding LexA-Aβ. Cells were grown in 100 ml of Glc or Gal/Raf medium (in which B42-Aβ expression was induced) to an OD_{600} of 0.6–0.8, pelleted by centrifugation, resuspended in 500 μl of RIPA buffer (75), lysed by beating with glass beads five times for 2 min each, and spun twice for 5 min in a microcentrifuge (10,000 × *g*) at 4°C to remove the beads and cell debris. Five microliters of the supernatant was taken as a control, and 15 μl of rabbit anti-LexA antiserum [kindly donated by Roger Brent (33)] was added to the remainder, which was incubated at 4°C for 4 hr on a rotating platform. LexA-containing proteins were precipitated from this remainder with 50 μl of protein A-Sepharose CL-4B (Sigma). The entire pellet was dissolved in Laemmli sample buffer, subjected to sodium dodecyl sulfate/polyacrylamide gel electrophoresis (SDS/PAGE; Integrated Separation System, Hyde Park, MA), and blotted onto nitrocellulose. Tagged Aβ fusion proteins were identified by Western analysis of the blotted proteins with the 12CA5 monoclonal anti-HA antibody (34). Cell extracts and immunoprecipitates were also subjected to immunoblotting with monoclonal anti-Aβ antibodies 4G8 and 6E10. Western blot analysis was performed with ECL chemiluminescence reagents using the protocol supplied by the vendor (Amersham).

Electron Microscopy. Dilutions of Aβ for incubation with the octapeptide QKLVTTAE were performed as in ref. 41 (ratio of Aβ to octapeptide was 1:10). The photomicrographs were obtained with a JEOL JEM-100S electron microscope at 80 kV (×155,000 magnification).

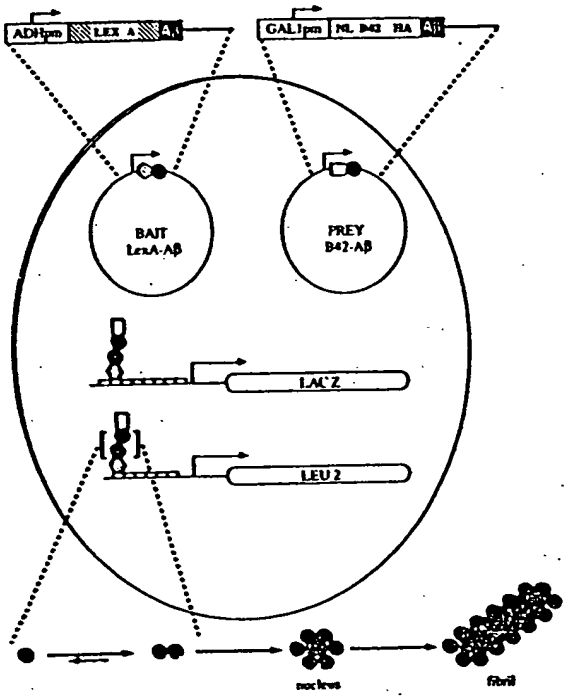


FIG. 1. Schematic representation of yeast strain EGY48 transformed with bait (LexA-Aβ fusion), prey (B42-Aβ fusion), and *lacZ* reporter plasmids. Bait fusion protein (LexA-Aβ) is produced constitutively under the control of the *ADH* promoter and binds to the upstream region of reporter genes at LexA operator sites; prey fusion peptide expression is driven by the *GAL1* promoter and is inducible by galactose. Prey fusion protein is expressed in the presence of galactose. If the Aβ portion of the prey protein binds the Aβ moiety of the bait fusion peptide, transcription from the reporter genes is triggered. The system therefore provides an opportunity to examine interaction between two monomeric Aβ molecules, an essential first step in the nucleation event leading to fibril formation.

BEST AVAILABLE COPY

RESULTS

The experimental system established by Brent *et al.* (described in ref. 29) is depicted in Fig. 1. The selection strain contains either Aβ or bicoid as a bait fused in-frame to the bacterial LexA protein, which by itself has no transcriptional activation function in yeast (33). The host strain contains *LEU2* and *lacZ* reporters carrying LexA operators instead of native upstream activating sequences. A strain containing the bait (LexA-Aβ) and reporters (*LEU2* and *lacZ*) remains inert for the expression of leucine utilization or β-galactosidase activity unless it also contains a vector (prey) that expresses an interacting protein as a fusion molecule consisting of nuclear localization sequences from simian virus 40, the B42 acid blob transactivation domain, and an epitope tag from influenza virus HA protein (35). In this system, conditional expression of library-encoded proteins is directed by the *GAL1* promoter (achieved by growing yeast cells in Gal/Raf minimal medium).

We first determined whether EGY48 strains containing the LexA protein alone, LexA-Aβ fusion protein, or LexA-bicoid permitted the expression of *lacZ* or leucine genes. When EGY48 strains containing the individual LexA fusion baits were spotted at equal density on minimal medium plates containing Gal/Raf Ura$^-$ His$^-$ medium, similar growth rates were observed, indicating that none of the baits was toxic to yeast. These strains failed to grow on Gal/Raf Ura$^-$ His$^-$ Leu$^-$ medium and did not form blue colonies on Gal/Raf Ura$^-$ His$^-$ X-Gal medium (data not presented), indicating that none of the bait proteins by themselves could permit the expression of leucine or β-galactosidase phenotypes.

B42-Aβ prey plasmid was introduced into the yeast strain containing LexA-Aβ bait protein. Equal dilutions of this yeast strain were spotted on Gal/Raf Ura$^-$ His$^-$ Trp$^-$ X-Gal and Glc Ura$^-$ His$^-$ Trp$^-$ X-Gal media to measure expression of β-galactosidase, and on Gal/Raf Ura$^-$ His$^-$ Trp$^-$ Leu$^-$ and Glc Ura$^-$ His$^-$ Trp$^-$ Leu$^-$ media to check the expression of the leucine utilization phenotype. The B42-Aβ prey plasmid, when introduced into the yeast strain with LexA-Aβ bait, showed growth on minimal medium plates devoid of leucine (Fig. 2*A*) and showed blue colonies on X-Gal medium in the presence of Gal/Raf as the carbon source (Fig. 2*C*) but showed no growth (Fig. 2*B*) and no blue colonies (Fig. 2*D*) in the presence of glucose. These results indicate that the interaction between LexA-Aβ and B42-Aβ is triggered by expression of the B42 fusion protein under the influence of the *GAL1* promoter. When LexA-bicoid (Fig. 2) or LexA protein alone (data not presented) was used as bait, introduction of B42-Aβ plasmid did not result in growth on leucine plates or blue colonies on X-Gal medium in the presence of Gal/Raf, indicating that the

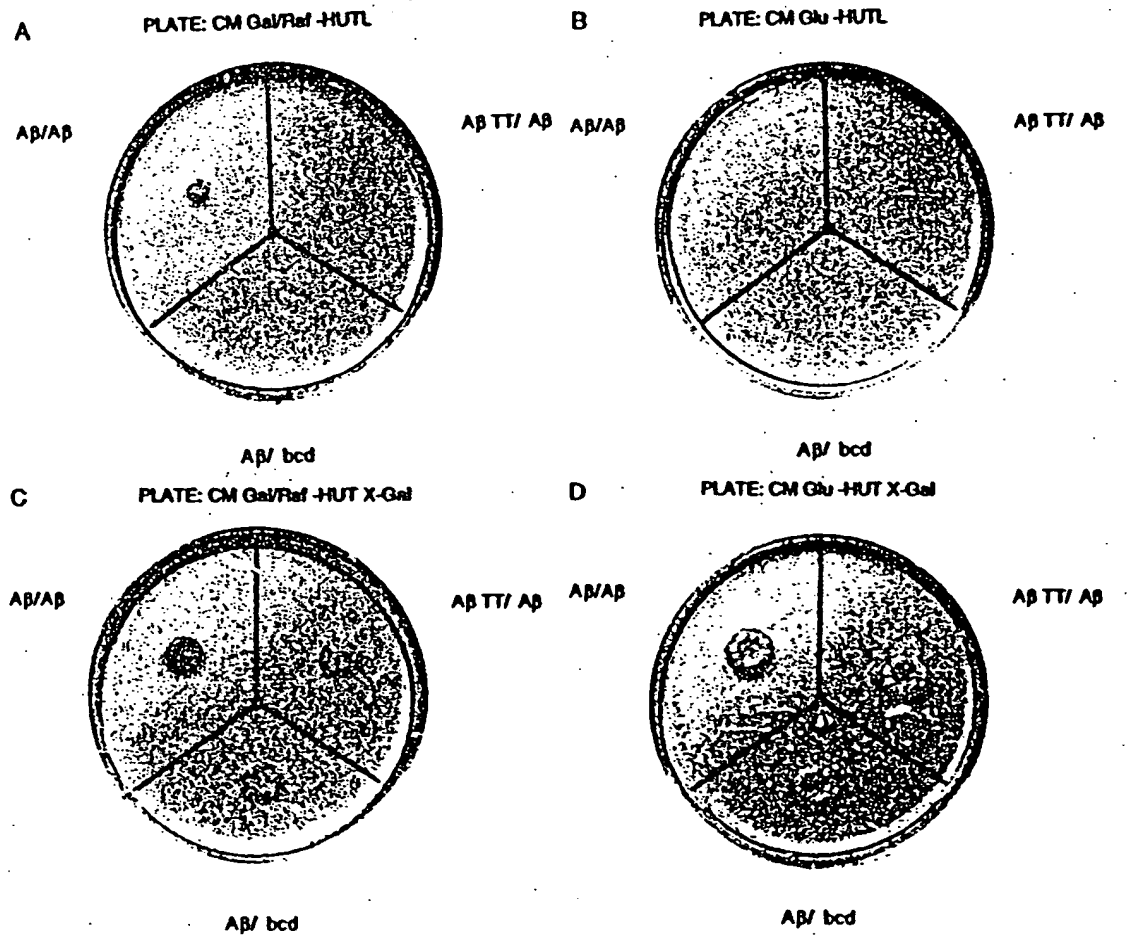


FIG. 2. Four-plate screen to examine interaction. B42-Aβ/LexA-Aβ strain, 342-AβTT/LexA-Aβ strain, and B42-Aβ/LexA-bcd strain were grown on plates containing Gal/Raf Ura$^-$ His$^-$ Trp$^-$ Leu$^-$ minimal medium (*A*), Glc Ura$^-$ His$^-$ Trp$^-$ Leu$^-$ minimal medium (*B*), Gal/Raf Ura$^-$ His$^-$ Trp$^-$ X-Gal minimal medium (*C*), or Glc Ura$^-$ His$^-$ Trp$^-$ X-Gal minimal medium (*D*).

BEST AVAILABLE COPY

interaction between LexA-Aβ and B42-Aβ is specific and most likely due to the intermolecular interaction between the Aβ molecules derived from the bait and the prey. When B42-AβTT prey plasmid was introduced into the yeast strain bearing LexA-Aβ bait, minimal growth was observed in plates devoid of leucine and no blue (or only very weakly blue) colonies were observed on X-Gal medium in the presence of glucose or Gal/Raf as carbon source (Fig. 2). This indicates that the Aβ molecule substituted at positions 19 and 20 with threonine residues interacts poorly with the wild-type Aβ peptide.

We next attempted to quantitate the observed Aβ–Aβ interaction by the ONPG colorimetric assay. Fig. 3 clearly indicates that there was significantly higher β-galactosidase activity in the yeast strain expressing B42-Aβ prey and LexA-Aβ bait compared with the yeast cells with B42-Aβ prey/LexA-bicoid bait (~2.5-fold, $P = 0.01$, Student's t test) or B42-AβTT prey/LexA-Aβ bait (~2-fold, $P = 0.02$). These results indicate that the Aβ–Aβ interaction inferred from Fig. 2 was significantly greater than the interaction between AβTT and Aβ or Aβ and bicoid monomers.

In an attempt to obtain direct *in vivo* evidence for the interaction between B42-Aβ prey and LexA-Aβ bait proteins, immunoprecipitates obtained by using antibodies against bait protein were subjected to immunoblotting with antibodies against the prey protein. Yeast cells expressing LexA-Aβ bait and B42-Aβ or B42-AβTT prey proteins were grown in glucose-containing medium and switched to glucose or Gal/Raf liquid minimal medium. The cells were harvested and cell extracts were prepared from equal numbers of cells. One aliquot of cell extract was subjected to immunoprecipitation with an anti-LexA antibody and the immunoprecipitates were subjected to Tris/tricine SDS/PAGE followed by immunoblotting with the monoclonal anti-HA antibody. If the two Aβ molecules or Aβ–AβTT molecules interact *in vivo*, one should be able to isolate the bait–prey complexes with antibody specific to the bait. Indeed, prey-specific HA immunoreactivity was observed (at 16 kDa; Fig. 4, lane 1) from the immunoprecipitates obtained from Aβ/Aβ cells grown in the presence of galactose, but not from the immunoprecipitates obtained from these cells subjected to glucose in the medium (Fig. 4, lane 2), indicating that the two Aβ fusion proteins interact inside the yeast cell nucleus. When Aβ/Aβ cell extracts were directly subjected to immunoblotting with anti-HA antibody, the 16-kDa band was observed in extracts derived from these cells grown in the presence of galactose (Fig. 4, lane 3), but no

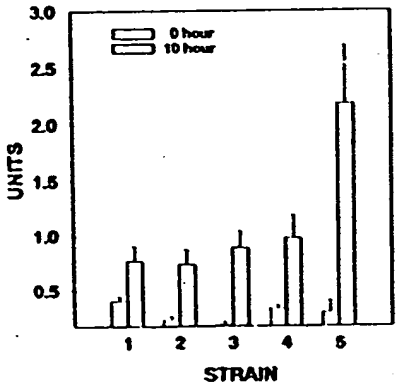


FIG. 3. Assay to determine β-galactosidase activity present in each of the strains tested after 0 hr (open bars) and 10 hr (filled bars) of incubation in Gal/Raf complete minimal medium. Bars: 1, B42-Aβ/LexA-bicoid; 2, B42 alone/LexA alone; 3, B42 alone/LexA-Aβ; 4, B42-AβTT/LexA-Aβ; 5, B42-Aβ/LexA-Aβ. These data are typical of three replicated experiments.

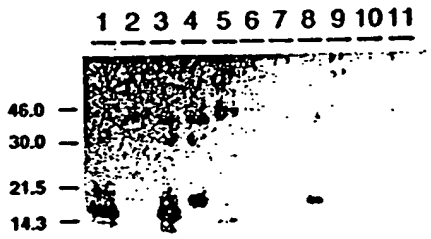


FIG. 4. Immunoprecipitates obtained by using LexA antiserum with extracts of B42-Aβ/LexA-Aβ strain grown in Gal/Raf medium (lane 1) or Glc medium (lane 2), cell extracts of B42-Aβ/LexA-Aβ cells grown in the presence of Gal/Raf (lane 3) or Glc (lane 4). Lanes 5–8 are similar to lanes 1–4 except that B42-AβTT/LexA-Aβ yeast strain was used. Lanes 9 and 10 represent immunoprecipitates and cell extracts obtained from LexA-Aβ strain containing no prey plasmid and from B42 alone/LexA alone yeast strain, respectively. Samples were electrophoresed in a Tris/tricine SDS/15% polyacrylamide gel and immunoblotted with anti-HA antibody.

immunoreactive band was observed for cells grown in the presence of glucose (Fig. 4, lane 4). The LexA immunoprecipitates obtained from AβTT/Aβ yeast grown on galactose resulted in very low levels of the anti-HA-immunoreactive 16-kDa band, seen only upon prolonged exposure (Fig. 4, lane 5; band not seen at this exposure). No 16-kDa band resulted from immunoprecipitates grown on glucose even on prolonged exposure (Fig. 4, lane 6). The anti-HA-immunoreactive 16-kDa band was, however, observed in cell extracts obtained from AβTT/Aβ yeast grown on galactose (Fig. 4, lane 7) but not in AβTT/Aβ yeast grown on glucose (Fig. 4, lane 8; the AβTT prey protein seems to run slightly lower than the Aβ prey protein). The observed molecular mass of 16 kDa is consistent with that predicted for a prey fusion protein and was also observed when anti-Aβ antibodies were used with cell extracts from Aβ/Aβ cells grown on galactose (data not presented). No immunoreactive bands were obtained in LexA immunoprecipitates derived from the EGY48 strain expressing the LexA-Aβ bait but containing no prey plasmid (Fig. 4, lane 9) and in cell extracts or immunoprecipitates from E9Y48 strain.

The observed weak interaction of AβTT with the native Aβ molecule was also examined by electron microscopy. An octapeptide fragment, QKLVTTAE, representing the FF-to-TT substitution at positions 19 and 20 in Aβ-(17–24) is capable of inhibiting fibril formation of the native Aβ-(1–40) (Fig. 5). Aβ-(1–40) at 250 μM, incubated for 4 days in water, showed significantly greater fibril formation (Fig. 5*A*) compared with the amount seen when Aβ-(1–40) at 250 μM was incubated with the octapeptide at 2.5 mM under the same conditions (Fig. 5*B*).

DISCUSSION

In the present study we demonstrate that two monomers of Aβ are capable of interacting in a eukaryotic cell. We further demonstrate that this interaction is specific by using the *Drosophila* protein bicoid as a bait protein in this system. This interaction was found to be positive by using the *lacZ* and *LEU2* reporter systems (Fig. 2). Quantitation of β-galactosidase activity by the ONPG assay supported these conclusions (Fig. 3). Furthermore, direct evidence of interaction was obtained by subjecting immunoprecipitates obtained by using antibodies against bait protein to immunoblotting with antibodies raised against the HA epitope present on the prey protein (Fig. 4).

Hilbich *et al.* (36) have previously reported that a well-preserved hydrophobic core around aa 17–24 is important for

BEST AVAILABLE COPY

the formation of β-sheet structure and amyloid properties. When stained with Congo red, peptide Aβ-(10–42) or Aβ-(10–43) containing the FF-to-TT substitution (equivalent to the substitution in AβTT) did not reveal birefringence and showed decreased β-sheet content when compared with the native peptide by circular dichroism spectroscopy and by infrared spectroscopy. Moreover, when these substituted peptides were mixed with the native Aβ-(10–43) fragment at equimolar concentration, they inhibited the formation of filaments *in vitro* (by ≈ 15%), as detected by electron microscopic analysis (36). Our data suggest that the octapeptide fragment QKLVTTAE, representing the FF-to-TT substitution in Aβ-(17–24), is also capable of inhibiting the aggregation of native Aβ-(1–40) (Fig. 5; unpublished results). The results presented in this paper clearly demonstrate that AβTT fusion protein interacts poorly with the native Aβ (Figs. 2–4). Hilbich *et al.* (36) suggest that the Aβ and AβTT monomers may form oligomers that do not fit into the structure of a filament. Our data suggest that inhibition of filament formation by peptides representing the TT substitution (at 10-fold molar abundance compared with the native peptide) may be explained by a weak interaction between the Aβ and AβTT monomers. Alternatively, it is also possible that AβTT peptide may directly interfere with the fiber formation process.

Recent evidence has indicated that the cellular forms of prion protein (PrP^c) can form protease-resistant prion protein (PrP^{res}) in a cell-free system in which PrP^c is used as a seed (37). This conversion did not require biosynthesis of new PrP^c, asparagine-linked glycosylation, or the presence of its normal glycosylphosphatidylinositol anchor, suggesting that oligomer

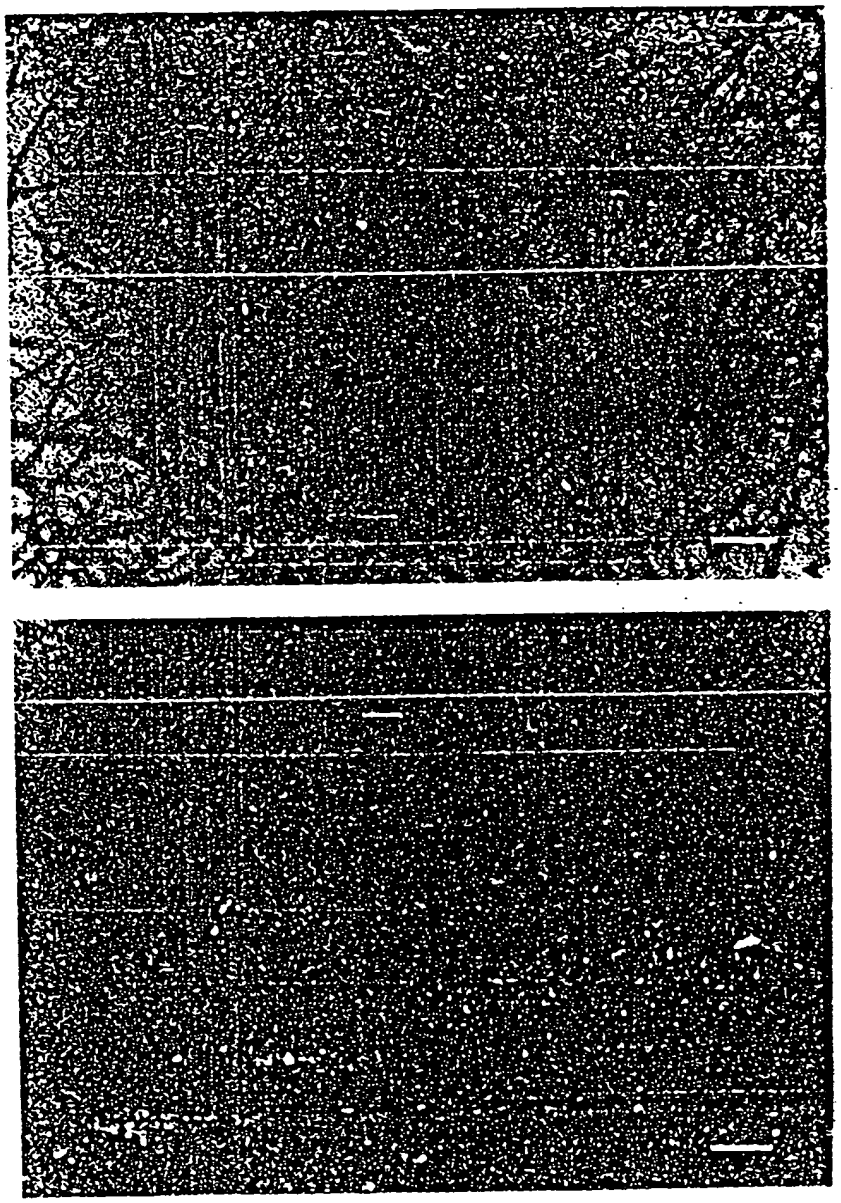


FIG. 5. Transmission electron micrographs of Aβ-(1–40) peptide incubated for 4 days in water (*A*) and Aβ-(1–40) peptide incubated in a 1:10 ratio with the octapeptide OKLVTTAE for 4 days in water (*B*). (Bar = 0.1 μm.)

BEST AVAILABLE COPY

formation between PrP^{sc} and PrP^{c} is sufficient for the conversion reaction to occur. Amyloid fibrils characterize several human diseases in which the presence of amyloid deposits is coincident with organ dysfunction. There is often a positive correlation between severity of the disease and the extent of fibril formation (38). Amyloid formation exhibits nucleation-dependent kinetics (22, 39–41). The slow and thermodynamically unfavorable interactions between individual monomers may be the rate-limiting step in aggregation. The yeast system described in this paper offers an opportunity to study the interaction of monomeric Aβ peptides. Although the peptides were expressed as fusion proteins, the results presented in this paper suggest that the observed interaction is due to the Aβ peptide domain. Results presented in Fig. 4 also suggest that no covalent higher-order bait–prey aggregates can be observed on the gel. This system may therefore provide an opportunity to freeze-frame the monomer–monomer interaction. Experimental system(s) such as the one presented in this paper may thus be used to study monomer–monomer interactions in other proteins that participate in nucleation-dependent amyloid formation.

There are some caveats in the present study. Although our results indicate that the interaction of LexA-Aβ and B42-Aβ fusion proteins occurs mainly because of amino acids within the two Aβ domains, it is possible that the conformation of the fusion proteins may influence this effect. These interactions take place in the yeast cell nucleus under conditions where the LexA-Aβ bait complex is bound to the LexA operator site. Whereas the yeast system described here is useful in studying intermolecular interactions, the intramolecular interactions may not be fully captured in a fusion-protein context. Furthermore, it is possible that interaction between bait-Aβ and prey-Aβ may not be relevant to fibril formation. This system will therefore have to be carefully validated by using molecules that are known to accelerate or inhibit the monomer–monomer interaction in Aβ fibrillogenesis.

We thank Dr. Roger Brent (Massachusetts General Hospital, Boston) for providing the necessary vectors and yeast strains required for completion of this study. We are very thankful to Dr. Timm Jessen for help with several experimental protocols, Paula Williams for electron microscopy, and Michael Merriman for oligonucleotide and peptide synthesis. The monoclonal antibodies to Aβ peptide were provided by Drs. K. S. Kim and H. Wisniewski.

1. Masters, C. L., Simms, G., Weinman, N. A., Multhaup, G., McDonald, B. L. & Beyreuther, K. (1985) *Proc. Natl. Acad. Sci. USA* 82, 4245–4249.
2. Glenner, G. G. & Wong, C. W. (1984) *Biochem. Biophys. Res. Commun.* 122, 1131–1135.
3. Kang, J., Lemaire, H. G., Unterbeck, A., Salbaum, J. M., Masters, C. L., Grzeschik, K. H., Multhaup, G., Beyreuther, K. & Muller-Hill, B. (1987) *Nature (London)* 325, 733–736.
4. Tanzi, R. E., Gusella, J. F., Watkins, P. C., Bruns, G. A., St. George-Hyslop, P., Van Keuren, M. L., Patterson, D., Pagan, S., Kurnit, D. M. & Neve, R. L. (1987) *Science* 235, 880–884.
5. Goldgaber, D., Lerman, M. I., McBride, O. W., Saffiotti, U. & Gajdusek, D. C. (1987) *Science* 235, 877–880.
6. Robakis, N. K., Ramakrishna, N., Wolfe, G. & Wisniewski, H. M. (1987) *Proc. Natl. Acad. Sci. USA* 84, 4190–4194.
7. Levy, E., Carman, M., Fernandez-Madrid, I., Power, M., Lieberburg, I., van Duinen, S., Gerard, T., Bots, A., Luyendijk, W. & Frangione, B. (1990) *Science* 248, 1124–1128.
8. Goate, A., Chartie-Harlin, M. C., Mullan, M., Brown, J., Crawford, F., *et al.* (1991) *Nature (London)* 349, 704–706.
9. Murrell, J., Farlow, M., Ghetti, B. & Benson, M. D. (1991) *Science* 254, 97–99.
10. Mullan, M., Crawford, F., Axelman, K., Houlden, H., Lilius, L., Winblad, B. & Lannfelt, L. (1992) *Nat. Genet.* 1, 345–347.
11. Jones, C. T., Morris, S., Yates, C. M., Moffoot, A., Sharpe, C., Brock, D. J. & St. Clair, D. (1992) *Nat. Genet.* 1, 306–309.
12. Hendriks, L., van Duijn, C. M., Cras, P., Cruts, M., Van Hul, W., van Harskamp, F., Warren, A., McInnis, M. G., Antonarakis, S. E., Martin, J. J., Hofman, A. & Van Broeckhaven, C. (1992) *Nat. Genet.* 1, 218–221.
13. Citron, M., Oltersdorf, T., Haass, C., McConlogue, L., Hung, A. Y., Seubert, P., Vigo-Pelfrey, C., Lieberburg, I. & Selkoe, D. J. (1992) *Nature (London)* 360, 672–674.
14. Cai, X. D., Golde, T. E. & Younkin, S. G. (1993) *Science* 259, 514–516.
15. Shoji, M., Golde, T. E., Ghiso, J., Cheung, T. T., Estus, S., Shaffer, L. M., Cai, X.-D., McKay, D. M., Tintner, R., Frangione, B. & Younkin, S. G. (1992) *Science* 258, 126–129.
16. Seubert, P., Vigo-Pelfrey, C., Esch, F., Lee, M., Dovey, H., Davis, D., Sinha, S., Schlossmacher, M., Whaley, J., Swindlehurst, C., McCormack, R., Wolfert, R., Selkoe, D., Lieberburg, I. & Shenk, D. (1992) *Nature (London)* 359, 325–327.
17. Haass, C., Schlossmacher, M. G., Hung, A. Y., Vigo-Pelfrey, C., Mellon, A., Ostaszewski, B. L., Lieberburg, I., Koo, E. H., Schenk, D., Teplow, D. B. & Selkoe, D. J. (1992) *Nature (London)* 359, 322–325.
18. van Gool, W. A., Schenk, D. B. & Bolhuis, P. A. (1994) *Neurosci. Lett.* 172, 122–124.
19. Miller, D. L., Papayannopoulos, I. A., Styles, J., Bobin, S. A., Lin, Y. Y., Biemann, K. & Iqbal, K. (1993) *Arch. Biochem. Biophys.* 301, 41–52.
20. Mori, H., Takio, K., Ogawara, M. & Selkoe, D. J. (1992) *J. Biol. Chem.* 267, 17082–17086.
21. Iwatsubo, T., Odaka, A., Suzuki, N., Mizusawa, H., Nukina, N. & Ihara, Y. (1994) *Neuron* 13, 45–53.
22. Jarrett, J. T., Berger, E. P. & Lansbury, P. T., Jr. (1993) *Biochemistry* 32, 4693–4697.
23. Jarrett, J. T. & Lansbury, P. T., Jr. (1993) *Cell* 73, 1055–1058.
24. Kelly, J. W. & Lansbury, P. T., Jr. (1994) *Amyloid Int. J. Exp. Clin. Invest.* 1, 186–205.
25. Pike, C. J., Walencewicz-Wasserman, J., Kosmoski, J., Cribbs, D. H., Glabe, C. G. & Cotman, C. W. (1995) *J. Neurochem.* 64, 253–265.
26. Ma, J., Yee, A., Brewer, H., Jr., Das, S. & Potter, H. (1994) *Nature (London)* 372, 92–94.
27. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. & Struhl, K., eds. (1994) *Current Protocols in Molecular Biology* (Wiley, New York), pp. 13.6.2–13.7.5.
28. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY).
29. Golemis, E. A. & Brent, R. (1992) *Mol. Cell Biol.* 12, 3006–3014.
30. Ma, J. & Ptashne, M. (1987) *Cell* 51, 113–119.
31. Gill, G. & Ptashne, M. (1988) *Nature (London)* 334, 721–724.
32. Harlow, E. & Lane, D. (1988) *Antibodies: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY), 2nd Ed.
33. Brent, R. & Ptashne, M. (1984) *Nature (London)* 312, 612–615.
34. Samson, J.-L., Jacson-Grusby, L. & Brent, R. (1989) *Cell* 57, 1045–1052.
35. Gyuris, J., Golemis, E., Chertkov, H. & Brent, R. (1993) *Cell* 75, 791–803.
36. Hilbich, C., Kisters Woike, B., Reed, J., Masters, C. L. & Beyreuther, K. (1992) *J. Mol. Biol.* 228, 460–473.
37. Kocisco, D. A., Corne, J. H., Priola, S. A., Cheseborn, B., Raymond, G. J., Lansbury, P. T. & Caughey, B. (1994) *Nature (London)* 370, 471–474.
38. Hawkins, P. N., Cavender, J. P & Pepys, M. B. (1990) *N. Engl. J. Med.* 323, 508–513.
39. Colon, W. & Kelly, J. W. (1992) *Biochemistry* 31, 8654–8660.
40. Jarrett, J. T. & Lansbury, P. T., Jr. (1992) *Biochemistry* 31, 12345–12352.
41. Corne, J. H., Fraser, P. E. & Lansbury, P. T., Jr. (1993) *Proc. Natl. Acad. Sci. USA* 90, 5959–5963.

MICROBIOLOGICAL REVIEWS, Mar. 1995, p. 94–123
0146-0749/95/$04.00+0
Copyright © 1995, American Society for Microbiology
Vol. 59, No. 1

BEST AVAILABLE COPY

Protein-Protein Interactions: Methods for Detection and Analysis

ERIC M. PHIZICKY[1]* AND STANLEY FIELDS[2]

Department of Biochemistry, University of Rochester Medical School, Rochester, New York 14642,[1] and Department of Molecular Genetics and Microbiology, State University of New York at Stony Brook, Stony Brook, New York 11794[2]

INTRODUCTION 95
PHYSICAL METHODS TO SELECT AND DETECT PROTEINS THAT BIND ANOTHER PROTEIN 96
Protein Affinity Chromatography 96
Purity of the coupled protein and use of protein fusions 97
Influence of modification state 97
Retention of native structure of the coupled protein 97
Concentration of the coupled protein 98
Amount of extract applied 98
Other considerations 98
Affinity Blotting 98
Immunoprecipitation 99
Cross-Linking 100
Determination of architecture 100
Detection of interacting proteins 101
(i) Detection in vivo 101
(ii) Detection in vitro 101
(iii) Other considerations 102
LIBRARY-BASED METHODS 102
Protein Probing 102
Phage Display 103
Basic approach 103
Related methods 104
(i) Antibody phage 104
(ii) Peptides on plasmids 105
Two-Hybrid System 105
Other Library-Based Methods 106
GENETIC METHODS 106
Extragenic Suppressors 107
Synthetic Lethal Effects 107
Overproduction Phenotypes 108
Overproduction of wild-type proteins 108
Overproduction of mutant proteins 108
Unlinked Noncomplementation 109
POPULAR METHODS TO ESTIMATE AND DETERMINE BINDING CONSTANTS 109
Importance of Characterization of the Binding Interaction 109
Binding constant 109
Concentrations of species 109
Influence of competing proteins 109
Influence of cofactors 109
Effect of cellular compartmentation 109
Solution conditions 109
Limits of Binding-Constant Considerations 109
Methods for Determining Binding Constants 110
Binding to immobilized proteins 110
Sedimentation through gradients 111
Gel filtration columns 111
(i) Nonequilibrium "small-zone" gel filtration columns 111
(ii) Hummel-Dreyer method of equilibrium gel filtration 111
(iii) Large-zone equilibrium gel filtration 112
Sedimentati n equilibrium 112

* Corresponding author. Mailing address: Department of Biochemistry, University of Rochester Medical School, 601 Elmwood Ave., Box 607, Rochester, NY 14642. Phone: (716) 275-7268. Fax: (716) 271-2683.

Flu rescence methods 112
(i) Fluorescence spectrum 113
(ii) Fluorescence polarizati n or anis tropy with tagged molecules 113
Solution equilibrium measured with immobilized binding protein 114
Surface plasmon resonance 114
Limits to Detection 115
EXAMPLES OF WELL-CHARACTERIZED DOMAINS 115
Leucine Zipper 116
Structure 116
Stability 116
Specificity 116
Regulation 116
SH2 Domain 116
SH3 Domain 117
CONCLUDING REMARKS 118
ACKNOWLEDGMENTS 118
REFERENCES 118

INTRODUCTION

Protein-protein interactions are intrinsic to virtually every cellular process. Any listing of major research topics in biology—for example, DNA replication, transcription, translation, icing, secretion, cell cycle control, signal transduction, and ermediary metabolism—is also a listing of processes in which protein complexes have been implicated as essential components. In consequence, the analysis of the proteins in ese complexes is no longer the exclusive domain of biochemss; geneticists, cell biologists, developmental biologists, molecular biologists, and biophysicists have by necessity all gotten into the act. We attempt in this review to summarize both ssical and recent methods to identify proteins that interact d to assess the strengths of these interactions.

Proteins that are composed of more than one subunit are found in many different classes of proteins. Some of the best-aracterized multisubunit proteins are those that, as originally rified, contained two or more different components. These include classical proteins such as hemoglobin, tryptophan synthetase, aspartate transcarbamylase, core RNA polymerase, -replicase, and glycyl-tRNA synthetase. Since these prons purified as multisubunit complexes, their protein-protein interactions were self-evident.

Other well-known examples of multisubunit proteins include ch more complicated assemblies of polypeptides. These inde metabolic enzymes such as the pyruvate dehydrogenase and α-ketoglutarate dehydrogenase complexes, the DNA replication complex of *Escherichia coli* and other organisms, the cterial flagellar apparatus, the nuclear pore complex, and the l assembly of bacteriophage T4. Also included in this group are ribonucleoprotein complexes, such as the signal recognition particle of the glycosylation pathway, small nuclear ribocleoproteins of the spliceosome, and the ribosome itself. hough some of the subunits of these protein complexes are not tightly bound, activity is associated with a large structure that in many cases is called a protein machine (5).

Ther are also a large number of transient protein-protein eractions, which in turn control a large number of cellular processes. All modifications of proteins necessarily involve such transient protein-protein interactions. These include the eractions of protein kinases, protein phosphatases, glycosyl nsferases, acyl transferases, proteases, etc., with their substrate proteins. Such protein-modifying enzymes encompass a large number of protein-protein interactions in the cell and ulate all manner of fundamental processes such as cell wth, cell cycle, metabolic pathways, and signal transduction.

Transient protein-protein interactions are also involved in the recruitment and assembly of the transcription complex to specific promoters, the transport of proteins across membranes, the folding of native proteins catalyzed by chaperonins, individual steps of the translation cycle, and the breakdown and re-formation of subcellular structures during the cell cycle (such as the cytoplasmic microtubules, the spindle apparatus, nuclear lamina, and the nuclear pore complex). Transient complexes are much more difficult to study, because the proteins or conditions responsible for the transient reaction have to be identified first. Part of the goal of this review is to describe recent methods and developments that have allowed their identification and characterization.

Protein-protein interactions can have a number of different measurable effects. First, they can alter the kinetic properties of proteins. This can be reflected in altered binding of substrates, altered catalysis, or (as first enunciated by Monod et al. [153]) altered allosteric properties of the complex. Thus, the interaction of proliferating-cell nuclear antigen with DNA polymerase δ alters the processivity of the polymerase (174), the interaction of succinate thiokinase and α-ketoglutarate dehydrogenase lowers the K_m for succinyl coenzyme A by 30-fold (171), and the cooperative binding of oxygen to hemoglobin and the allosteric regulation of aspartate transcarbamylase are regulated by interactions of the protomers. Second, protein-protein interactions are one common mechanism to allow for substrate channeling. The paradigm for this type of complex is tryptophan synthetase from *Neurospora crassa*. It is a complex of two subunits, each of which carries out one of the two steps of reaction (formation of indole from indole 3-glycerol phosphate, followed by conversion of indole to tryptophan). The intermediate indole is noncovalently bound, but it is preferentially channeled to form tryptophan (241). Many similar examples of metabolic channeling have been demonstrated, both between different subunits of a complex and between different domains of a single multifunctional polypeptide (see reference 208 for a review). Third, protein-protein interactions can result in the formation of a new binding site. Thus, an ADP site forms at the interface of the α and β subunits of *Escherichia coli* F_1-ATPase (228), yeast hexokinase binds one ATP molecule at the interface of the asymmetric homodimer (209), and phosphofructokinase from *Bacillus stearothermophilus* binds both fructose 6-phosphate and ADP at the interface between subunits (60). Fourth, protein-protein interactions can inactivate a protein; this is the case with the interaction of phage P22 repressor with its antirepressor (213), with the interaction of trypsin with trypsin inhibitor (221), and with th interaction of

BEST AVAILABLE COPY

phage T7 gene 1.2 protein with *E. coli* dGTP triphosphohydrolase (156). Fifth, protein-protein interactions can change the specificity of a protein for its substrate; thus, the interaction of lactalbumin with lactose synthase lowers the K_m for glucose by 1,000-fold (95), and the interaction of transcription factors with RNA polymerase directs the polymerase to different promoters.

Klotz et al. (116) enumerated four advantages of multisubunit proteins relative to a single large protein with multiple sites. First, it is much more economical to build proteins from simpler subunits than to require multiple copies of the coding information to synthesize oligomers. Thus, for example, actin filaments and virus coats are much more simply assembled from monomers than by translation of a large polyprotein of repeated domains. Similarly, it is much more convenient to have one gene encoding a protein with different interacting partners, such as some of the eukaryotic RNA polymerase subunits, than to have the gene for that subunit reiterated for each different polymerase. Second, translation of large proteins can cause a significant increase in errors in translation; if such errors cause a lack of activity, they are much more economically eliminated by preventing assembly of that subunit into the complex than by eliminating the whole protein. Third, multisubunit assemblies allow for synthesis at one locale, followed by diffusion and assembly at another locale; this allows for both faster diffusion (since the monomers are smaller) and compartmentalization of activity (if assembly is required for activity). Fourth, homooligomeric proteins, if they have an advantage over monomers, are easily selected in evolution if the oligomers interact in an antiparallel arrangement; in this case, a single-amino-acid change that increases interaction potential has effects at two such sites.

Another advantage of multisubunit complexes is the ability to use different combinations of subunits to alter the magnitude or type of response. Thus, for example, adult hemoglobin ($\alpha 2\beta 2$) and fetal hemoglobin ($\alpha 2\gamma 2$) are each composed of heterooligomers with a common α subunit; differences in the binding of oxygen in these hemoglobins allow oxygen to be readily passed from mother to fetus. Other examples include the oligomerization of Jun with Fos or with itself, which results in distinct activities in transcription because the different dimers bend DNA in opposite directions (114); the interaction of TATA-binding protein with the transcription apparatus of RNA polymerase I, II, or III, in which TATA-binding protein plays different roles (235); the interactions of microtubules with the large set of proteins to which they bind (113), not all of which bind at the same time; the interaction of different transcription factors with core RNA polymerase in both eukaryotes and prokaryotes to direct transcription of different genes; and the interaction of retinoblastoma (Rb) protein with viral oncoproteins and other cellular proteins (31, 32).

Protein-protein interactions may be mediated at one extreme by a small region of one protein fitting into a cleft in another protein and at another extreme by two surfaces interacting over a large area. Examples of the first case include the large class of protein-protein interactions that involve a domain of a protein interacting tightly with a small peptide. The paradigm for this type of interaction is that of specific Src homology 2 (SH2) domains with specific small peptides containing a phosphotyrosyl residue. This interaction occurs with a dissociation constant as low as nM and is du to a specific binding pocket in SH2 domains not unlike a classical substrate-binding pocket (64, 205, 224, 225). Many other examples of domains that bind small peptides with affinities in the nanomolar to m lar range have been described. Th paradigm for the second case, i.e., surfaces that interact with each other over large areas, is that of the leucine zipper, in which a stretch of α-helix forms a surface that fits almost perfectly with another α-helix from another subunit protein (59, 161; also see reference 4). Binding also occurs in th nanomolar range for such interactions (196). Other interactions may occur through intermediate-sized complementary surfaces.

It is evident that protein-protein interactions are much more widespread than once suspected, and the degree of regulation that they confer is large. To properly understand their significance in the cell, one needs to identify the different interactions, understand the extent to which they take place in the cell, and determine the consequences of the interaction. This review is intended to supply an overview of three aspects of protein-protein interactions. First, we briefly describe a number of physical, molecular biological, and genetic approaches that have been used to detect protein-protein interactions. Second, we describe several experimental approaches that have been used to evaluate the strength of protein-protein interactions. Third, we describe three well-characterized domains that are responsible for protein-protein interactions in a number of different proteins. As the literature on this topic is vast, we have not attempted to conduct an exhaustive review. Rather, we hope that this article serves as a journeyman's guide to protein-protein interactions.

The first and still the most comprehensive review on protein-protein interactions is that of Klotz et al. (116). This review contains a survey of the subunit composition and binding energies of all oligomeric proteins that had been identified at the time, as well as a discussion of the geometry of interactions and an excellent discussion of the influence of binding constants, concentrations, and cooperativity parameters on th population of oligomers. A good discussion of channeling and compartmentation is found in the monograph by Friedrich on quaternary structure (70) and the article by Srere (208). The review by Eisenstein and Schachman (57) contains an interesting discussion of the functional roles of subunits of oligomeric proteins and of approaches used to determine whether the monomers of oligomeric proteins are active. Also of interest is the discussion of proteins as machines (5) and a discussion of protein size and composition (78).

PHYSICAL METHODS TO SELECT AND DETECT PROTEINS THAT BIND ANOTHER PROTEIN

Protein Affinity Chromatography

A protein can be covalently coupled to a matrix such as Sepharose under controlled conditions and used to select ligand proteins that bind and are retained from an appropriate extract. Most proteins pass through such columns or are readily washed off under low-salt conditions; proteins that are retained can then be eluted by high-salt solutions, cofactors, chaotropic solvents, or sodium dodecyl sulfate (SDS) (Fig. 1). If the extract is labeled in vivo before the experiment, there are two distinct advantages: labeled proteins can be detected with high sensitivity, and unlabeled polypeptides derived from the covalently bound protein can be ignored (these might be either proteolytic fragments of the covalently bound protein or subunits of the protein which are not themselves covalently bound). This method was first used 20 years ago to detect phage and host proteins that interacted with different forms of *E. coli* RNA polymerase (177). Proteins that were retained by an RNA polymerase-agarose column (which was shown to be enzymatically active) but not by a control column coupled with bovine serum albumin were judged as interacting candidates. The interactions were substantiated in two ways. First, the

BEST AVAILABLE COPY

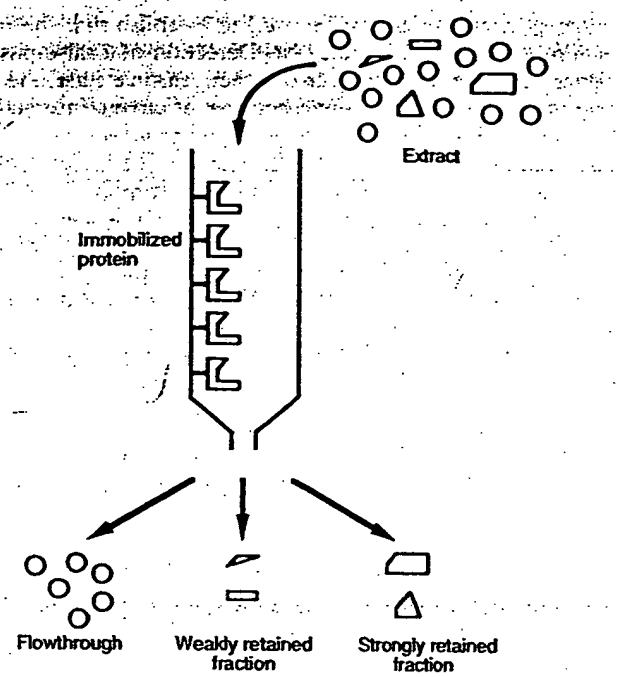


FIG. 1. Protein affinity chromatography. Extract proteins are passed over a column containing immobilized protein. Proteins that do not bind flow through the column, and ligand proteins that bind are retained. Strongly retained proteins have more contacts with the immobilized protein than do those that are weakly retained.

interaction of T7 0.3 protein with RNA polymerase was confirmed by coimmunoprecipitation of the 0.3 protein with RNA polymerase antibody. Second, the interaction of T4 proteins with RNA polymerase was shown to depend on the form of RNA polymerase on the column: one T4 protein interacted with core RNA polymerase and T4-modified RNA polymerase but not with RNA polymerase holoenzyme, and another interacted only with the T4-modified polymerase. The phage proteins that bound RNA polymerase were identified by their absence in appropriate T4 and T7 mutants.

Similar methods have been used, particularly by the laboratories of J. Greenblatt and B. Alberts, to identify many other protein-protein interactions. Two excellent reviews on the topic, which cover many of the details of coupling and a number of strategic considerations, have been published (69, 145).

Candidat proteins can be coupled directly to commercially available preactivated resins as described by Formosa et al. (69). Alternatively, they can be tethered noncovalently through high-affinity binding interactions. Thus, Beeckmans and Kanarek (14) demonstrated an interaction between fumarase and malate dehydrogenase by immobilizing the test enzyme with antibody bound to protein A-Sepharose, as well as by direct covalent coupling of the test enzyme to Sepharose. Some of the important considerations of a successful binding experiment are elaborated below.

Purity of the coupled protein and use of protein fusions. An essential requirement for a successful protein affinity chromatography experiment is pure prot in; otherwise, any interacting protein that is detected might be binding to a contaminant in the preparation. Greenblatt and Li (80) did two experiments to establish that core RNA polymerase bound to NusA on the column rather than to a contaminant in the NusA preparation. First, they demonstrated that a fully active NusA variant protein, which presumably contained different amounts of various contaminants (since it eluted at different positions in columns used to purify it), still bound cor RNA polymerase; second, they demonstrated by independent experiments that th complex contained equimolar amounts of NusA protein and core RNA polymerase.

The easiest way to obtain pure protein, if the gene is available, is through the use of protein fusions. Several such systems have been described; in each case, the protein of interest (or a domain of the protein) is fused to a protein or a domain that can be rapidly purified on the appropriate affinity resin. The most common such fusion contains glutathione *S*-transferase (GST), which can be purified on glutathione-agarose columns (202). Other fusions in common use include *Staphylococcus* protein A, which can be purified on columns bearing immunoglobulin G; oligohistidine-containing peptides, which can be purified on columns bearing Ni^{2+}; the maltose-binding protein, which can be purified on resins containing amylose; and dihydrofolate reductase, which can be purified on methotrexate columns. (Other common protein fusions which add an epitope for the influenza virus hemagglutinin [12CA5] or c-Myc are also in common use and are used most often for coimmunoprecipitation [see the section on immunoprecipitation, below].)

Purified fusion proteins are used in two ways to detect interactions on affinity columns. First, the protein is covalently coupled to the resins in the usual way, as was done by Mayer et al. (139) to detect a tyrosine-phosphorylated protein that bound to the SH2 domain of Abl tyrosine kinase and by Weng et al. (232) to demonstrate that the SH3 domain of c-Src binds paxillin. Second, the purified fusion proteins can be noncovalently bound to the beads and then mixed with an appropriate extract or protein. This was done by Zhang et al. (248) to demonstrate an interaction of the N-terminal portion of c-Raf with Ras, by Flynn et al. (68) to detect the binding of an actin filament-associated protein to Src-SH3/SH2, and by Hu et al. (99) to demonstrate the binding of the SH2 domain of the p85 subunit of phosphatidylinositol 3-kinase to two different growth factor receptors.

Influence of modification state. The interactions of many proteins with their target proteins often depends on the modification state of one or both of the proteins (mostly by phosphorylation). Thus, the recognition of Rb protein by the transcription factor E2F and by the transforming proteins simian virus 40 large T antigen, human papillomavirus-16 E7, and adenovirus E1A is more efficient with underphosphorylated than phosphorylated Rb (132, 133, 240). Conversely, SH2 domains of proteins, for example, recognize tyrosine phosphorylated substrates several orders of magnitude more efficiently than they do their nonphosphorylated counterparts (64). Protein-protein interactions that require a posttranslationally modified protein for interaction are not detected if the protein is purified by the use of expression vectors in cells in which the protein is not properly modified. A means to circumvent this problem is to use GST fusion vectors to express proteins in host cells more related to their origin. Thus, the interaction of bovine papillomavirus E5 oncoprotein with an α-adaptin-like molecule was confirmed by addition of beads to extracts of NIH 3T3 cells that were expressing the GST-E5 fusion (38). Similarly, a yeast GST vector that allows regulated expression of yeast GST fusion proteins has been described (148).

Retenti n of native structure of the coupled protein. Failure to detect an interacting protein can result from inactivation of the protein during coupling. Ideally, coupling would immobilize a protein or a complex by randomly tethering it to th matrix through one covalent bond. For example, binding of *E. coli* proteins to immobilized λ N protein occurred only when

BEST AVAILABLE COPY

the cyanogen bromide (CNBr)-activated residues on the matrix were partially inactivated before coupling; this was attributed to the large number of lysine residues in λ N protein and the generation of multiple (and denaturing) covalent bonds between λ N protein and the matrix if th concentration of CNBr-activated matrix sites was too high (80). Therefore, determining that the coupled protein has retained its native structure is an important control, when possible. With some proteins, such as RNA polymerase from *E. coli*, activity could be detected when the coupled protein was assayed on the matrix (177). With others, such as filamentous actin (F-actin) columns, the desired polymerized form was stabilized with phalloidin (or by chemical cross-linking), and the proteins that bound F-actin were shown not to bind monomeric actin (14). Similarly, microtubule columns were stabilized with taxol (113).

Native protein structure also depends on all subunits of a complex being present in the coupled resin. This can be assessed by SDS elution of a sample of the resin and comparison of the subunit composition of the eluted material with that of the starting material. In the case of *E. coli* RNA polymerase, all the components of the enzyme were still present (177). In the case of mammalian RNA polymerase II, one of the subunits did not reproducibly remain after coupling (206).

Concentration of the coupled protein. To detect interactions efficiently, the concentration of protein covalently bound to the column has to be well above the K_d of the interaction. Thus, for the detection of weak protein-protein interactions, the concentration of bound protein should be as high as possible. Weak interactions can be completely missed on columns with lower concentrations of coupled protein, even if they contain correspondingly larger amounts of resin to maintain the same total amount of bound protein (see the sections on importance of characterization of the binding interaction and on binding to immobilized proteins, below, for a discussion of this point).

Amount of extract applied. The amount of extract applied to the column can be critical for two opposing reasons. If too little extract is applied and the protein that binds is present at low concentration, too little protein will be retained to be detected, even if it binds with high affinity and is labeled with ^{35}S (see, for example, reference 206). Conversely, if too much protein is applied, competition among potential ligands may result in failure to detect minor species. This was observed by Miller and Alberts (144) in looking for minor protein species that interact with F-actin.

Other considerations. There are four distinct advantages of protein affinity chromatography as a technique for detecting protein-protein interactions. First, and most important, protein affinity chromatography is incredibly sensitive. With appropriate use (high concentrations of immobilized test protein), it can detect interactions with a binding constant as weak as 10^{-5} M (69) (see the section on binding to immobilized protein, below). This limit is within range of the weakest interaction likely to be physiologically relevant, which we estimate to be in the range of 10^{-3} M (see the section on limits of binding-constant considerations, below). Second, this technique tests all proteins in an extract equally; thus, extract proteins that are detected have successfully competed for the test protein with the rest of the population of proteins. Third, it is easy to examine both th domains of a protein and the critical residues within it that are responsible for a specific interaction, by preparing mutant derivatives (38, 216). Fourth, interactions that depend on a multisubunit tethered protein can be detected, unlike the case with protein blotting.

One potential problem derives from the very sensitivity of the technique. Sinc it detects interactions that are so weak, independent criteria must be used to establish that th interaction is physiologically relevant. Detection of a false-positive signal can arise for a number of other reasons. First, the protein may bind the test protein because of charge interactions; for this reason, it is desirabl to use a control column with approximately the same ionic charges. Second, the proteins may interact through a second protein that interacts with th test protein; although interesting in itself, the interaction may not be direct. Third, the proteins may interact with high specificity even though they never encounter one another in the cell. The most famous example of this type is the high affinity of actin for DNase I (125).

For all of these reasons, the prudent course is to independently demonstrate the interaction in vitro or, if possible, in vivo. Cosedimentation was used to confirm the interaction of RAP 72 (now known as RAP 74) and RAP 30 with RNA polymerase II (206), NusA protein with core RNA polymerase (80), and NusB protein with ribosomal protein S10 (138). In other cases, more biological criteria were used. For example, antibodies were generated against many of the proteins that interacted with F-actin (but not monomeric G-actin) on columns, and these were used to demonstrate that more than 90% of the corresponding proteins were localized with an actin-like distribution during mitosis of *Drosophila* embryos at th syncytial blastoderm stage of development (144). The identification of three yeast actin-binding proteins was confirmed in three separate ways: one of the proteins was shown to correspond to the yeast analog of myosin by virtue of a shared epitope; another protein colocalized with actin cables and cortical actin patches, and overproduction of the third protein caused a reorganization of the actin cytoskeleton (53). In the identification of microtubule-associated proteins, two criteria were used to demonstrate the authenticity of the results (113). First, antibodies for 20 of the 24 candidate microtubule-associated proteins stain various parts of microtubule structures of *Drosophila* embryos during the cell cycle. Second, many (but not all) of the microtubule-associated proteins isolated on microtubule affinity columns are the same as those isolated by traditional cosedimentation methods of Vallee and Collins (219).

Failure to detect an interaction can occur for a number of technical reasons, described above. A false-negative result can arise for two additional reasons: the interacting protein may not be able to exchange with another protein to which it is binding, or the two proteins may not be able to interact both with each other and with the resin.

Protein affinity chromatography does not always yield answers corresponding to other approaches. For reasons that are unclear, a large number of proteins were detected by probing SDS-polyacrylamide gel electrophoresis (PAGE) gels with a GST fusion of the SH2 domain of Abl tyrosine kinase, but only a couple of proteins were detected on columns coupled with this protein (139). Similarly, a specific protein was detected on F-actin columns stabilized by suberimidate cross-linking but not with phalloidin (144). Finally, G-actin interacting proteins are very difficult to detect with columns of G-actin, although such columns bind DNase I; by contrast, DNase I columns can be used to detect such G-actin interactions (24).

Affinity Blotting

In a procedure analogous to the use of affinity columns, proteins can be fractionated by PAGE transferred t a nitrocellulose membrane, and identified by their ability to bind a protein, peptide, or other ligand. This method is similar to immunoblotting (Western blotting), which uses an antibody as

BEST AVAILABLE COPY

the probe. Complex mixtures of proteins, such as total-cell lysates, can be analyzed without any purification. Therefore, this method has been particularly useful for membran proteins, such as cell surface receptors (see reference 207 for a discussion). Cell lysates can also be fractionated before gel electrophoresis to increase the sensitivity of the method for detecting interaction with rare proteins.

Considerations in affinity blotting include the biological activity of the proteins on the membrane, the preparation of the protein probe, and the method of detection. Denaturing gels, which are run in the presence of SDS and sulfhydryl reducing agents, will inactivate most proteins and separate subunits of a complex. These denaturants are removed during the blotting procedure, which allows many proteins to recover (or partially recover) activity. However, if biological activity is not recoverable, the proteins can be fractionated by a nondenaturing gel system. This method eliminates the problem of regeneration of activity and allows the detection of binding in cases when binding requires the presence of a protein complex.

The protein probe can be prepared by any one of several procedures, and, as with affinity columns, the recent advent of using tags to the protein has greatly facilitated this purification. Synthesis in *E. coli* with a GST fusion, epitope tag, or other affinity tag is most commonly used. The protein of interest can then be radioactively labeled, biotinylated, or used in the blotting procedure as an unlabeled probe that is detected by a specific antibody. Vectors that incorporate into the protein a short amino acid sequence recognized by the heart muscle cyclic AMP (cAMP)-dependent protein kinase provide another convenient means for in vitro labeling (18).

One example of affinity blotting is the study of calmodulin-binding proteins (77). Calmodulin can be ^{125}I labeled and used either to probe a gel strip directly or to probe a nitrocellulose membrane after transfer of fractionated proteins. Because the extent of renaturation of calmodulin-binding proteins is variable, the assay is not quantitative. False-positive results have been detected in which a basic sequence binds calmodulin, although generally this binding is Ca^{2+} independent. A major advantage of this technique is that in the analysis of a multimeric protein that binds calmodulin, the precise binding polypeptide can be readily identified by affinity blotting with calmodulin. Using a combination of genetic approaches, Geiser et al. (73) identified the spindle pole body component Spc110 (Nuf1) as interacting with yeast calmodulin and then used affinity blotting to demonstrate that labeled calmodulin could directly detect a GST-Spc110 fusion transferred to a blot after fractionation by SDS-PAGE.

Affinity blotting has been widely used in studies of the association of the regulatory subunit of the type II cAMP-dependent protein kinase with numerous specific anchoring proteins (reviewed in reference 29). Two-dimensional procedures of isoelectric focusing followed by SDS-PAGE have been used to increase the separation of these anchoring proteins. As a control in some of these experiments, a mutant of the regulatory subunit that is deleted for the first 23 residues did not detect any anchoring proteins.

Immunoprecipitation

Coimmunoprecipitation is a classical method of detecting protein-protein interactions and has been used in literally thousands of experiments. The basic experiment is simple. Cell lysates are generated, antibody is added, the antigen is precipitated and washed, and bound proteins are eluted and analyzed. Several sources of material are in wide use. The antigen used to make the antibody can be purified protein (either from the natural tissue or organism or purified after expression in another organism) or synthetic peptide coupled to carrier, and the antibody can be polyclonal or monoclonal. Alternatively, the protein can carry an epitope tag for which commercially available antibodies are available (12CA5 and c-Myc are in common use) or a protein tag (such as GST) for which beads are available to rapidly purify the GST fusion protein and any copurifying proteins. Glutathione-agarose beads were used, for example, to detect and characterize a GTP-dependent interaction of Ras and Raf (227) and to demonstrate that the v-Crk SH2 domain binds the phosphorylated form of paxillin (16). The GST fusion immunoprecipitates a 70-kDa protein that reacts with anti-paxillin antibody and with anti-phosphotyrosine antibody; moreover, anti-paxillin immunoprecipitates a protein that reacts with anti-Crk antibody but only under conditions when the paxillin is phosphorylated.

Several criteria are used to substantiate the authenticity of a coimmunoprecipitation experiment. First, it has to be established that the coprecipitated protein is precipitated by the antibody itself and not by a contaminating antibody in the preparation. This problem is avoided by the use of monoclonal antibodies. Polyclonal antibodies are usually preadsorbed against extracts lacking the protein to remove contaminants or are prepurified with authentic antigen. Peptide-derived antisera (which are usually made by coupling of the peptide to a carrier protein) are usually preadsorbed against the protein that was coupled, to remove antibody against the carrier, in addition to the usual purification to remove contaminating antibody. Second, it has to be established that the antibody does not itself recognize the coprecipitated protein. This can be accomplished by demonstrating persistence of coprecipitation with independently derived antibodies, ideally with specificities toward different parts of the protein. Alternatively, it can sometimes be demonstrated that coprecipitation requires the presence of the antigen; cell lines, growth conditions, or strains that lack the protein cannot coprecipitate the protein unless the antigen is added. In certain cases, it can also be shown that antibody generated against the coprecipitated protein will coprecipitate the original antigen. Third, one would like to determine if the interaction is direct or proceds through another protein that contacts both the antigen and the coprecipitated protein. This is usually addressed with purified proteins, by immunological or other techniques. Fourth, and most difficult, is determining that the interaction takes place in the cell and not as a consequence of cell lysis. Such proteins ought to colocalize, or mutants ought to affect the same process.

A particularly good example of this technique is the demonstration that adenovirus E1A protein interacts with Rb protein. A mixture of monoclonal antibodies against E1A coimmunoprecipitated a discrete set of five polypeptides (and some smaller ones) from a cell line expressing E1A, including a particularly abundant one of 110 kDa (84). Four lines of evidence supported the claim that the 110-kDa polypeptide was forming a complex with E1A protein. First, coprecipitation was not specific to a single antibody; three independent monoclonal antibodies against E1A protein coimmunoprecipitated this protein. Second, these antibodies did not themselves recognize or immunoprecipitate the native or denatured 110-kDa protein, although they recognized and immunoprecipitated native and denatured E1A protein. Third, coprecipitation required E1A protein; the 110-kDa polypeptide could be immunoprecipitated from HeLa extracts (which do not contain E1A protein) only if a source of E1A protein was added. Fourth, the complex could be detected independently in crude lysates; a subpopulation of E1A protein in lysed cells sedimented at 10S rather than at 4S, and this subpopulation contained coimmu-

BEST AVAILABLE COPY

noprecipitable 110-kDa protein. A similar 110-kDa protein (as well as a similar set of other proteins) was also identified with antipeptide antisera against E1A protein (242). Two separat antisera (one against an amino-terminal peptide and one against a carboxyl-terminal peptide) each coprecipitated the 110-kDa polypeptide, and coprecipitation was prevented either with an excess of the corresponding E1A peptide antigen or in cell extracts lacking E1A protein.

Subsequent studies established that this 105- to 110-kDa polypeptide was the Rb gene product (236). To this end, monoclonal antibodies against the 110-kDa protein were prepared by immune purification of the 110-kDa protein. The resulting antibody coprecipitated E1A protein, just as anti-E1A coprecipitated the 110-kDa protein. Since the 110-kDa protein was the same size as Rb protein, and since it was present in a wide variety of cell lines but not in cell lines known to contain deletions of the Rb gene, it seemed likely that the 110-kDa protein was Rb protein. This was proved by using anti-Rb peptide antibodies against different regions of Rb in three experiments. First, 110-kDa protein precipitated with anti-110-kDa antibody comigrated and had the same partial peptide map as that precipitated with anti-Rb antibody. Second, 110-kDa protein precipitated with anti-E1A antibody could be detected in immunoblots with two different anti-Rb antibodies, and this detection was inhibited by the corresponding peptide antigen. Third, anti-110-kDa antibody could immunoprecipitate Rb protein synthesized in vitro.

When coimmunoprecipitation is performed with unsupplemented crude lysates, as is often the case, this technique has four distinct advantages. First, like protein affinity chromatography, it detects the interactions in the midst of all the competing proteins present in a crude lysate; therefore, the results from this sort of experiment have a built-in specificity control. Second, both the antigen and the interacting proteins are present in the same relative concentrations as found in the cell; therefore, any artificial effects of deliberate overproduction of the test protein are avoided. Third, elaborate complexes are already in their natural state and can be readily coprecipitated; such complexes might otherwise be difficult to assemble in vitro. Fourth, the proteins are present in their natural state of posttranslational modification; therefore, interactions that require phosphorylation (or lack of phosphorylation) are more realistically assessed. Two disadvantages are also apparent. First, coimmunoprecipitating proteins do not necessarily interact directly, since they can be part of larger complexes. For example, the coprecipitation of E1A and p60 (now known to be cyclin A) (84) occurs indirectly; E1A interacts with p107 (237), and p107 interacts with cyclin A (61, 62). Similarly, coprecipitation of Rb protein with E2F probably occurs through another protein (92, 179). Second, coprecipitation is not as sensitive as other methods, such as protein affinity chromatography, because the concentration of the antigen is lower than it is in protein affinity chromatography. This can be overcome by deliberately adding an excess of the antigen to the crude lysates to drive complex formation, as was done to detect a 46-kDa protein that competed with simian virus 40 T antigen for Rb protein (100). It can also be overcome by covalently cross-linking the proteins prior to immunoprecipitation (48) see the section on cross-linking, below). These alterations of course perturb the natural conditions that make immunoprecipitation an attractive method.

Cross-Linking

Cross-linking is used in two ways to deduce protein-protein interactions. First, it is used t deduce the architecture of proteins or assemblies that are readily isolated intact from the cell. Second, it is used to detect proteins that interact with a given test protein ligand by probing extracts, whole cells, or partially purified preparations.

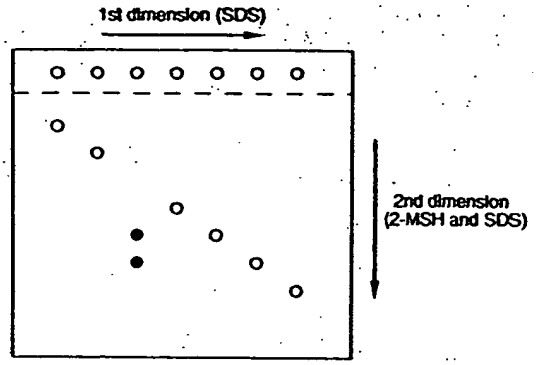


FIG. 2. Two-dimensional gels to identify cross-linked proteins in a complex. Proteins that are not cross-linked have the same mobility in both dimensions of the SDS gel and form a diagonal. Proteins that are cross-linked migrate slowly in the first dimension; after cleavage of the cross-link with mercaptoethanol (2-MSH), these proteins migrate at their native positions in the second dimension and are off the diagonal.

Determination of architecture. The classical method of identifying interacting partners in a purified protein complex involves analysis by two-dimensional gels (Fig. 2). The procedure involves three steps. First, the complex is reacted with a cleavable bifunctional reagent of the form RSSR', and the R and R' groups react with susceptible amino acid side chains in the protein complex. This reaction forms adducts of the form P-RSSR'-P'. Second, the proteins are fractionated on an SDS-gel in the absence of reducing agents. The gel separates the proteins based on molecular weight, and cross-linked proteins of the form P-RSSR'-P' migrate as species of greater molecular weight. Third, a second dimension of the SDS-gel is run after treatment of the gel with a reducing agent to cleave th central S—S bond. Un-cross-linked species align along the diagonal, because their molecular weights do not change after reduction. Cross-linked proteins migrate off the diagonal because they migrated as P-RSSR'-P' in the first dimension and as molecules of the form P-RSH and P'-R'SH in the second dimension. The cross-links are identified by their size, which corresponds to that of the un-cross-linked species P and P'. This method has been discussed at a practical step-by-step level by Traut et al. (215).

Cross-linking has been used to study the architecture of multienzyme complexes such as CF_1-ATPase (7) and *E. coli* F_1-ATPase (21). It has also been used to study the structure of much more complicated structures like the ribosome (41, 215). Since these structures are complex, the corresponding cross-linking pattern is necessarily complex. Furthermore, as might be expected, different patterns are sometimes obtained as the reactive group is changed and as the distance between the reactive groups is altered (41, 215). Several approaches have been taken to simplify the cross-linking patterns resulting from these experiments. In one approach, the proteins are prefractionated on urea-acrylamide gels or on CM-Sepharose before diagonal electrophoresis (41, 217). A second approach involves running two-dimensional gels without cleaving the cross-link, followed by elution of individual species, cleavage of th cross-link, and resolution of th resulting proteins on a third gel (22). A third approach involves th use of antibody to identify cross-linked partners after the use of appropriate gels (180, 212). Transfer of th gels followed by immunoblotting allows one to

BEST AVAILABLE COPY

unequivocally identify each member of a cross-linked pair. Since this method is so powerful, one-dimensional gels often suffice and noncleavable cross-linking reagents are easily used. Since immunoblotting is also very sensitive, one can take care to limit cross-linking to acceptably low levels.

Detection of interacting proteins. (i) Detection in vivo. Cross-linking in vivo can be accomplished with membrane-permeable cross-linking reagents followed by immunoprecipitation of the ligand protein. This method was used to detect a 60-kDa protein that interacts with Ras (48). Immunoprecipitation of this protein required both immune sera and cross-linking and was inhibited when excess Ras was added before immunoprecipitation. Since the cross-linked protein could be released from the immune complex by cleavage of the cross-link with dithiothreitol (but not by incubation of the immune complex in buffer), it was truly cross-linked. Since pretreatment of the cross-linking reagent with excess amino groups inhibited cross-linking but excess amino groups did not inhibit cross-linking if cells were lysed in their presence, cross-linking must have occurred in vivo. The complex was reproducibly increased after mitogenic stimulation and could be detected in cells producing normal amounts of Ras. This experiment makes another point: at least in these experiments, cross-linking before immunoprecipitation is a more sensitive technique than immunoprecipitation alone.

(ii) Detection in vitro. The addition of an isolated protein or peptide to a complex system offers a huge potential for precise and powerful cross-linking methods. Several different such methods have been used to detect interacting proteins.

(a) Labeled peptide or protein. Detection of cross-linking partners is incomparably cleaner if the protein or peptide is labeled before cross-linking, because there is only one source of labeled material. For example, ^{125}I-labeled gamma interferon was used to detect receptors that were cross-linked (192), and in vivo labeled interleukin-5 was purified before cross-linking to detect interacting receptors (147).

Proteins are also routinely labeled in vitro with [^{35}S]methionine during translation, and this was followed by cross-linking and by immunoprecipitation to detect protein interactions. This has been done, for example, to detect interaction of preprolactin and pre-β-lactamase with signal sequence receptor and translocation chain-associating protein during glycosylation (79) and to detect mitochondrial import proteins in contact with translocation intermediates (195, 204).

(b) Photoaffinity cross-linking with labeled cross-linking reagent. A particularly useful reagent is the Denny-Jaffee reagent, a cleavable heterobifunctional photoactivatable cross-linking reagent that is labeled on the photoactivated moiety (49). This reagent can be coupled to an isolated protein, which is then incubated in an appropriate extract and photoactivated to cross-link nearby proteins. Since the label is on the photoactivatable moiety of the cross-linking reagent, it is transferred to the cross-linked protein after cleavage of the cross-linking reagent (Fig. 3). This cross-linking reagent has been used to identify a specific 56-kDa ZP3-binding protein on acrosome-intact mouse sperm (19). As much as 90% of the label initially on ZP3 could be transferred to the 56-kDa protein, and cross-linking was inhibited by excess unlabeled ZP3 protein. Moreover, ZP3 affinity columns retained a protein with the same molecular mass. This reagent has also been used to demonstrate that phospholamban interacts with a specific site on the ATPase from sarcoplasmic reticulum only when it is nonphosphorylated and the ATPase is in the Ca^{2+}-free state (106).

Another useful reagent of this type is ^{125}I-{S-[N-(3-iodo-4-azidosalicyl)cysteaminyl]-2-thiopyridine}, also called IAC, a cysteine-specific modifying reagent. This reagent was used to

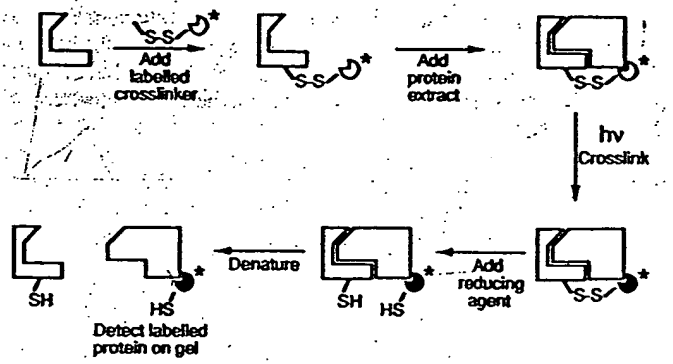


FIG. 3. Specific labeling of an interacting protein with a labeled photoactivatable cross-linking reagent.

demonstrate that the carboxy-terminal region of the subunit of *E. coli* RNA polymerase was adjacent to the activating domain of the catabolite activator protein (CAP) (33). To do this, a unique cysteine was introduced onto the surface of CAP, in a residue which tolerates a large number of mutations, and a preexisting surface cysteine was changed to serine. Subsequent reaction with labeled IAC resulted in quantitative incorporation of label and in protein with 70% of its transcription activation activity. Irradiation of the ternary complex of DNA, CAP, and RNA polymerase yielded 20% cross-linking, all of which was with a particular domain of the subunit of polymerase.

(c) Direct incorporation of photoreactive lysine derivative during translation. A photoactivatable group can be incorporated directly into the translation product by using a modified lysyl-tRNA. If translation is done in the presence of [^{35}S]methionine, the protein is simultaneously labeled and ready for photoactivated cross-linking. This approach has been particularly valuable in investigating the process by which proteins are inserted into the endoplasmic reticulum. During elongation, signal recognition particle (SRP) binds the nascent chain and halts translation until the arrested translation product is brought to the SRP receptor. This releases SRP, allowing translation to continue, coupled with translocation of the protein into the endoplasmic reticulum. With bovine preprolactin, there are two lysines at positions 4 and 9 of the signal sequence and no other lysine residues within the first 70 amino acids, after which translation is normally stopped by SRP. Thus, incorporation of lysine with a photoactivated group specifically probes interaction of the signal sequence with other interacting proteins. In this way, the nascent chain was specifically cross-linked with the 54-kDa protein of SRP and a 35-kDa microsomal membrane protein, called the signal sequence receptor (239). Subsequent experiments in the same system relied on translation of truncated mRNAs bearing lysine codons at different positions. These templates produce proteins that remain tethered to the ribosome through peptidyl-tRNA because of the lack of a termination codon. They therefore cannot complete translocation and are trapped, presumably as intermediates. In this way, it was shown that lysines in different positions also recognized the same 35-kDa membrane protein (121, 238). Moreover, this protein is probably required for translocation because antibodies against it inhibit translocation in vitro (87).

Investigation with the same system in *S. cerevisiae* demonstrated that prepro-α-factor is in contact with Sec61 protein (155). Antibody against either Sec61 or prepro-α-factor precipitated the same labeled cross-linked protein. Cross-linking

BEST AVAILABLE COPY

was observed only when prepro-α-factor was tethered; release of the protein with puromycin or a complete translation sequence abolished cross-linking. Moreover, the tethered prepro-α-factor was glycosylated while it was tethered, and cross-linking was ATP dependent for large tethered prepro-α-factor peptides; this indicated that prepro-α-factor had entered the normal glycosylation pathway. Sanders et al. (191) also demonstrated by conventional cross-linking followed by immunoprecipitation that Sec61 is in contact with tethered proteins being translocated (in this case by covalent coupling to avidin); the same experiments also demonstrated that BiP (Kar2) was cross-linked to the translocation intermediates and that *sec62* and *sec63* mutants modulate the process. The convergence of genetics and biochemical cross-linking studies further substantiates these interactions.

(d) Site-specific incorporation of photoreactive amino acid derivative during translation. Use of a suppressor tRNA to incorporate a photoactivatable amino acid derivative results in site-specific incorporation by use of a gene carrying a single stop codon. Two such reports have been described. High et al. (94) used a charged amber suppressor tRNA to insert a phenylalanine derivative into various regions of the signal sequence of preprolactin. Cross-linking experiments demonstrated that the amino-terminal end of the signal sequence is in proximity to the translocating chain-associating protein, whereas the hydrophobic core of the sequence contacts Sec61 protein. Cornish et al. (39) used a similar method to incorporate a different photoaffinity label. Still to be described is a similar method involving a labeled photoactivated amino acid replacement—the ultimate magic bullet.

(iii) Other considerations. One major disadvantage of using any cross-linking technique to detect protein-protein interactions is that it detects nearest neighbors which may not be in direct contact. The cross-linking reagent reaches out to any protein in close vicinity; thus, it may appear to detect protein interactions that are more like ships just passing in the night. This is more and more of a problem as the size of the cross-linking reagent increases. Any interaction detected by cross-linking should therefore be independently assessed for protein-protein interactions. However, cross-linking has three important advantages over other methods. First, it can "cement" weak interactions that would otherwise not be visible by other methods (see, for example, reference 48). Second, it can be used to detect transient contacts with different proteins at various stages in a dynamic process such as glycosylation, by freezing the process at different stages. Third, cross-linking can be done in vivo with membrane-permeable cross-linking reagents (48). It may also be possible to detect cross-linking in vivo after microinjection of a protein that is modified with a photoactivatable cross-linking group. To our knowledge, this has not yet been reported.

LIBRARY-BASED METHODS

A variety of methods have been developed to screen large libraries for genes or fragments of genes whose products may interact with a protein of interest. As these methods are by their nature highly qualitative, the interactions identified must be subsequently confirmed by biochemical approaches. However, the enormous advantag of these strategies is that the genes for these newly identified proteins or peptides are immediately availabl . This is in sharp contrast to th biochemical methods described in th section on physical methods to select and detect proteins that bind another protein, above, which generally result in the appearance of bands on a polyacrylamide gel. These library methods also differ from classical genetic techniques described in the section on genetic methods, below, which often require a specific phenotype before they can be carried out. Library screens are generally performed in bacteria or yeasts, organisms with rapid doubling times. Thus, these procedures can be completed rapidly.

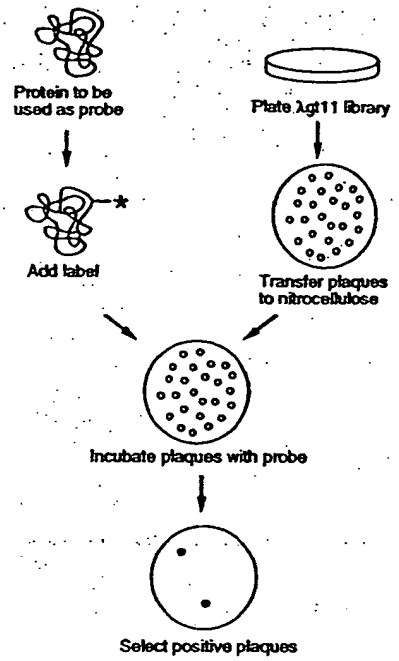


FIG. 4. Use of a labeled protein to probe an expression library.

Protein Probing

A labeled protein can be used as a probe to screen an expression library in order to identify genes encoding proteins that interact with this probe. Interactions occur on nitrocellulose filters between an immobilized protein (generally expressed in *E. coli* from a λgt11 cDNA library) and the labeled probe protein (Fig. 4). The method is highly general and therefore widely applicable, in that proteins as diverse as transcription factors and growth factor receptors have been used as probe. A variety of approaches can be used to label the protein ligand, or this ligand can be unlabeled and subsequently detected by specific antibody.

The method is based on the approach of Young and Davis (244), who showed that an antibody can be used to screen expression libraries to identify a gene encoding a protein antigen. The λgt11 libraries typically use an isopropyl-β-D-thiogalactopyranoside (IPTG)-inducible promoter to express proteins fused to β-galactosidase. Proteins from the bacteriophage plaques are transferred to nitrocellulose filters, incubated with antibody, and washed to remove nonspecifically bound antibody. Protein ligands were first used as probes in this type of experiment by Sikela and Hahn (200), who identified a brain calmodulin-binding protein with ^{125}I-labeled calmodulin as th probe. The λgt11-expressed fusion protein bound calmodulin with a K_d between 3 and 10 nM, and binding was dependent on the presence of Ca^{2+}. The signal-to-noise ratio in these experiments was higher than that found with various antibody probes.

MacGregor et al. (135) used the leucine zipper and DNA-binding domain of Jun as a probe and identified the rat cAMP response element-binding protein type 1. In this case, the Jun-

BEST AVAILABLE COPY

domain was biotinylated and detected with a streptavidin-alkaline phosphatase conjugate. Buffer conditions could be adjusted to distinguish a Jun-Jun homodimer from the more stable Fos-Jun heterodimer. Blackwood and Eisenman (17) used a similar approach with the basic-region helix-loop-helix leucine zipper domain (bHLH-zip) of the c-Myc protein. A 92-residue carboxy terminus of Myc, containing this domain, was expressed as a GST fusion protein, purified by glutathione-agarose affinity chromatography and ^{125}I labeled. This probe identified a new bHLH-zip protein termed Max, and gel shift experiments indicated that the Myc-Max complex exhibited site-specific DNA binding under conditions where neither Myc nor Max alone could bind. These results were critical in establishing a long-sought role for the Myc protein. Extending this result, Ayer et al. (6) used Max as a labeled probe to identify another member of this class, termed Mad.

A major advantage of the protein-probing approach is that the protein probe can be manipulated in vitro to provide, for example, a specific posttranslational modification or a metal cofactor. This modification or cofactor may be essential for the ability of the probe to bind to other proteins. This feature of the approach was exploited in the Ca^{2+}-dependent binding of calmodulin (200). Skolnik et al. (201) extended this use to phosphorylated probes in order to find proteins that bind to the carboxy-terminal phosphorylated tail of the epidermal growth factor (EGF) receptor. This tail is part of the intracellular domain of the receptor, which possesses a protein tyrosine kinase activity stimulated by binding of EGF. Skolnik et al. purified this domain from cells infected with a recombinant baculovirus, tyrosine phosphorylated it in vitro, and cleaved it to separate the phosphorylated carboxy-terminal tail from the protein kinase domain. Probing an expression library identified proteins containing the SH2 domain, which recognizes phosphotyrosyl-containing peptides. This cloning approach might be applied to the identification of proteins interacting with other activated phosphorylated receptors, including tyrosine- and serine-specific phosphatases as well as kinases. In addition, it should be possible to modify probe proteins by means other than phosphorylation to identify new proteins that recognize such modifications.

Probing expression libraries with labeled protein has numerous advantages. Since any protein or protein domain can be specifically labeled for use as a probe, the sophisticated arsenal of GST fusion vectors, other expression and tagging systems, and in vitro translation systems can be exploited; this makes preparation of the probe relatively straightforward. If specific antibody to the target protein is available, the probe protein need not be labeled; the antibody can be used in a second step to detect plaques that have bound the target protein. More than 10^6 plaques can be screened in an experiment, plating 5×10^4 plaques per 150-mm dish. The method not only results in the immediate availability of the cloned gene for the interacting protein but also can provide data regarding a specific domain involved in the interaction, because the λgt11 insert is often only a partial cDNA. Conditions of the wash cycles can be adjusted to vary the affinity required to yield a signal. As with many library-based methods, probing expression libraries compares equally all binary combinations of the probe protein and a library-encoded protein. Thus, less abundant proteins, proteins with weak binding constants, and proteins that temporally or spatially rarely interact with the probe protein in vivo can all be detected as long as their transcripts are present in the mRNA pool used to generat the library.

This method has certain intrinsic limitations. Proteins encoded by th library must be capable of folding correctly in *E. coli*, generally as fusion proteins, and of maintaining their structure on a nitrocellulose filter. However, proteins often can be renatured by subjecting the filters to a denaturation-renaturation cycle with 6 M guanidin hydrochloride as described by Vinson et al. (222). Binding conditions are arbitrarily imposed by the investigator, rather than reflecting the native environment of the cell. Since all combinations of protein-protein interactions are assayed, including those that might never occur in vivo, the possibility of identifying artifactual partners exists. In particular, the relative abundance of each potential partner expressed in a colony or plaque of the library is similar, instead of varying and potentially being compartmentalized as in the cell. Any posttranslational modifications necessary for efficient binding will generally not occur in bacteria (although some such modifications can be performed in vitro). Screening rather than direct selection is the means of detection, which inherently limits the number of plaques that can be assayed. The use of screening also restricts the further genetic manipulations that can be applied to the cDNA inserts. For example, in the analysis of point mutations, it is not possible to select directly for rare mutations that affect the interaction. Different protein probes are likely to behave variably in this approach, such that binding and washing conditions may have to be adjusted in each case to maximize the signal-to-noise ratio.

Phage Display

Basic approach. Smith (203) first demonstrated that an *E. coli* filamentous phage can express a fusion protein bearing a foreign peptide on its surface. These foreign amino acids were accessible to antibody, such that the "fusion phage" could be enriched over ordinary phage by immunoaffinity purification. Smith suggested that libraries of fusion phage might be constructed and screened to identify proteins that bind to a specific antibody. In the past few years, there have been numerous developments in this technology to make it applicable to a variety of protein-protein and protein-peptide interactions.

Filamentous phages such as M13, fd, and f1 have approximately five copies of the gene III coat protein on their surface; thus, a foreign DNA sequence inserted into this gene results in multiple copies of the fusion protein displayed by the phage. This is called polyvalent display. Similarly, the major coat protein encoded by gene VIII can also display a foreign insert (104). The gene VIII protein allows up to 2,700 copies of the insert per phage. Generally, polyvalent display is limited to small peptides (see the next section) because larger inserts interfere with the function of the coat proteins and the phage become poorly infective.

Random sequences can be inserted into gene III or gene VIII to generate a library of fusion phage (Fig. 5). Such a library can then be screened to identify specific phage that display any sequence for which there is a binding partner, such as an antibody. This screening is performed by a series of affinity purifications known as panning. The phage are bound to the antibody, which is immobilized on a plastic dish. Phage that do not bind are washed away, and bound phage are eluted and used to infect *E. coli*. Each cycle results in a 1,000-fold or greater enrichment of specific phage, such that after a few rounds, DNA sequencing of the tight-binding phage reveals only a small number of sequences. In addition to the advantage of high selectivity, a second advantage of this technology is that large phage libraries can be constructed (up to 10^9 to 10^{10} complexity) and the affinity purification step can be carried out at very high concentrations of phage ($>10^{13}$ phage per ml) (50). Third, the direct coupling of the fusion protein to its gene in a single phage allows the immediate availability of sequence

BEST AVAILABLE COPY

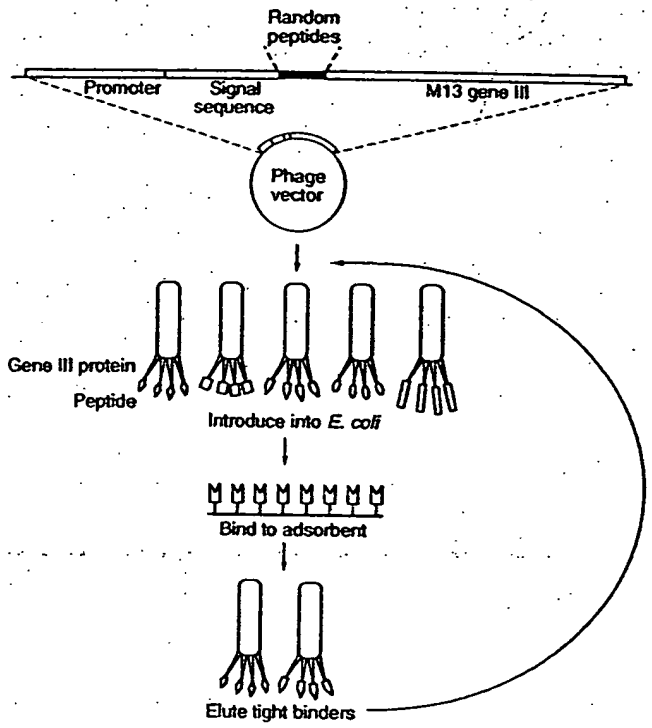


FIG. 5. A peptide library in a filamentous phage vector. The figure illustrates the process of panning, by which peptides that bind to an adsorbent are identified.

data to generate one or more consensus sequences of bound peptides or the sequences of variant proteins with a specific phenotype. Fourth, the phage can be used directly to assess the binding specificity of the encoded fusion proteins by varying the stringency of the wash procedures used in the panning cycles.

Random-sequence peptide libraries have been generated by cloning synthetic oligonucleotides into gene III (Fig. 5). Scott and Smith (198) generated a hexapeptide library and screened it to identify epitopes for two monoclonal antibodies specific for a hexapeptide from the protein myohemerythrin. Cwirla et al. (44) constructed a similar hexapeptide library to find peptides that can bind to a monoclonal antibody specific for a tetrapeptide from β-endorphin. Such epitope libraries allow rapid characterization of an unknown epitope recognized by either a monoclonal antibody or polyclonal serum. For example, monoclonal antibody pAB240, which recognizes the mutant conformation of the tumor suppressor p53 protein, was shown to bind to a 5-amino-acid motif in p53 (210). The binding partner for the phage-encoded peptides need not be an antibody. For example, Devlin et al. (50) constructed a 15-residue peptide library and used it to identify nine different peptides that bind to streptavidin.

A major advance in phage display came with the development of a monovalent system in which the coat protein fusion is expressed from a phagemid and a helper phage supplies a large excess of the wild-type coat protein (11, 131). Therefore, the phage are functional because the recombinant protein makes up only a small amount of th total coat protein. Th vast majority (>99%) of the population of phage particles display either on or no copies of the fusion protein on their surface. Such phage can accommodat 50 kDa of foreign protein without any significant effect on phage infectivity. In addition, monovalent phage display avoids potential avidity effects observed with polyvalent display, in which the phage can attach to the adsorbent at multiple points.

Phage display has also been used to identify proteins with increased binding affinity. In some cases, the use of monovalent display was necessary to avoid potential avidity effects, attributed to multipoint attachment of the polyvalent phage to the absorbent (231). Lowman et al. (131) expressed nearly on million mutants of human growth hormone (191 residues) as fusion phage and identified variants that bound tightly to the growth hormone receptor. The mutations were directed to 12 sites known to be important for binding to the receptor. Some variants had binding affinities up to eightfold greater than that of the wild-type hormone. Roberts et al. (186) used polyvalent display of bovine pancreatic trypsin inhibitor and directed mutagenesis to five residues of the protein. They selected for high-affinity inhibitors of human neutrophil elastase and identified one variant with an affinity 3.6×10^6 higher than that of wild-type bovine pancreatic trypsin inhibitor.

A similar strategy can be used with nontargeted mutagenesis. For example, Pannekoek et al. (167) expressed human plasminogen activator inhibitor 1, a 42-kDa protein, as a gene III protein fusion under conditions for monovalent display. The phage-displayed inhibitor could specifically form complexes with serine protease tissue-type plasminogen activator. PCR mutagenesis was used to generate a library of mutant plasminogen activator inhibitor 1 proteins, which can be screened to analyze structure-function relationships.

Phage display presents several advantages for the study of protein-protein interactions. The very large sizes of either random libraries or pools of individual variants of a single sequence that can be generated mean that complex mixtures can be screened. While not strictly a genetic approach, in that there is no direct selection for an interacting partner, phage display has many of the properties of genetic selection through its use of panning cycles. It is a rapid procedure and should be widely applicable. Although screening a random library of cDNA by a panning procedure to identify proteins that interact with a protein of interest has not yet been demonstrated, this strategy should prove workable.

Disadvantages of phage display include the size limitation of protein sequence for polyvalent display; the requirement for proteins to be secreted from *E. coli*; and the use of a bacterial host which may preclude the correct folding or modification of some proteins. All phage-encoded proteins are fusion proteins, which may limit the activity or accessibility for binding of some proteins. Since binding is detected in vitro, the sam considerations of an in vitro approach that are relevant for protein probing of expression libraries are relevant here.

Related methods. (i) Antibody phage. While we do not specifically address the vast topic of antigen-antibody interactions in this review, it is worth noting that phage display can be applied to these interactions. The principle of displaying antibody-combining domains on the surface of phage was first demonstrated by McCafferty et al. (141). The heavy- and light-chain variable domains of an anti-lysozyme antibody were linked on the same polypeptide and expressed as a gene III protein fusion. Over 1,000-fold enrichment of the antibody could be obtained by a single passage over a lysozym -Sepharose column. This method was then extended by this and other groups to allow the display of libraries of combining domains, such that new antibodies or mutant versions of existing antibodies could be generated.

Kang et al. (110) used a vector t express a combinatorial library of functional Fab molecules (~50-kDa heterodim r) on the surface of a phage. The Fd chain, consisting of the variable region and constant domain 1 of the immunoglobulin heavy

BEST AVAILABLE COPY

ain, was synthesized as a gene VIII protein fusion, while the ght chain contained no phage sequence. Th two chains d assembl in the bacterial periplasm and become incorated into the phag on coinfection with helper. Phage conined 1 to 24 antigen-binding sites per particle. The vector em described allows recombination of th two chains to erate large combinatorial libraries. A similar strategy to press Fabs by using the gene III protein has also been deibed (10). Additionally, a combinatorial library of linked vy- and light-chain variable genes fused to the gene III tein has been shown to be capable of detecting a highffinity binder (37). Kang et al. (110) suggested that such sysms can be used for mutation and selection cycles to generate -affinity antibodies. Moreover, they envisioned that the ems can be extended to analyze any protein recognition ystem, such as ligand-receptor interactions.

Phag display of Fab fragments was extended by Burton et (26), who generated a library of such fragments from the A of a human immunodeficiency virus-positive individual. After four rounds of panning with immobilized surface glycotein gp120 of the virus as the adsorbent, specific viral anodies were obtained. A similar method was used to obtain man antibody Fabs that recognize the hepatitis B surface antigen (246).

(iii) Peptides on plasmids. In a method highly analogous to ge display, random peptides are fused to the C terminus of *E. coli* Lac repressor and expressed from a plasmid that also contains Lac repressor-binding sites (43). Thus, the peptide fusions bind to the same plasmid that encodes them. The terial cells are lysed, and the peptide libraries are screened for peptides that bind to an immobilized receptor by using similar panning cycles to those for phage libraries. In this case, ptides become enriched because bound peptides carry their coding plasmids with them, via the repressor-operator interaction, and these plasmids can be transformed back into *E. coli*. In the initial example, peptides that bind to a monoclonal ibody specific for dynorphin B were selected, and these ptides contained a hexapeptide sequence similar to a segment of dynorphin B (43). This method is distinguished from the phage display methods in that the peptides are exposed at C terminus of the fusion protein and the fusions are cytosmic rather than exported to the periplasm.

Two-Hybrid System

The two-hybrid system (35, 65, 66) is a genetic method that es transcriptional activity as a measure of protein-protein interaction. It relies on the modular nature of many site-specic transcriptional activators, which consist of a DNA-binding main and a transcriptional activation domain (23, 97, 112). he DNA-binding domain serves to target the activator to the specific genes that will be expressed, and the activation domain ntacts other proteins of the transcriptional machinery to able transcription to occur. The two-hybrid system is based on the observation that the two domains of the activator need not be covalently linked and can be brought together by the eraction of any two proteins. The application of this system quires that two hybrids be constructed: a DNA-binding domain fused to some protein, X, and a transcription activation domain fused to some protein, Y. These two hybrids are expressed in a cell containing one or more reporter genes. If the and Y proteins interact, they create a functional activator by bringing the activation domain into close proximity with the DNA-binding domain; this can be detected by expression of the report r genes (Fig. 6). While th assay has been generally rformed in yeast cells, it works similarly in mammalian cells

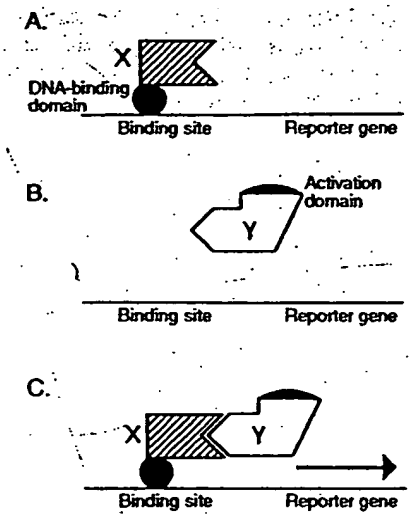


FIG. 6. The two-hybrid system. (A) The DNA-binding domain hybrid does not activate transcription if protein X does not contain an activation domain. (B) The activation domain hybrid does not activate transcription because it does not localize to the DNA-binding site. (C) Interaction between X and Y brings the activation domain into close proximity to the DNA-binding site and results in transcription.

(see, e.g., reference 46) and should be applicable to any other eukaryotic cells.

This method has been used with a wide variety of proteins, including some that normally reside in the nucleus, cytoplasm, or mitochondria, are peripherally associated with membranes, or are extracellular (see reference 66 for a review). It can be used to detect interactions between candidate proteins whose genes are available by constructing the appropriate hybrids and testing for reporter gene activity (220, 249). If an interaction is detected, deletions can be made in the DNA encoding one of the interacting proteins to identify a minimal domain for interaction (35). In addition, point mutations can be assayed to identify specific amino acid residues critical for the interaction (127). Most significantly, the two-hybrid system can be used to screen libraries of activation domain hybrids to identify proteins that bind to a protein of interest. These screens result in the immediate availability of the cloned gene for any new protein identified. In addition, since multiple clones that encode overlapping regions of protein are often identified, the minimal domain for interaction may be readily apparent from the initial screen (105, 223).

A variety of versions of the two-hybrid system exist, commonly involving DNA-binding domains that derive from the yeast Gal4 protein (35, 55) or the *E. coli* LexA protein (223, 247). Transcriptional activation domains are commonly derived from the Gal4 protein (35, 55) or the herpes simplex virus VP16 protein (45). Reporter genes include the *E. coli lacZ* gene (65) and selectable yeast genes such as *HIS3* (55) and *LEU2* (247). An increasing number of activation domain libraries are becoming available, such that screens are now feasible for proteins from many different organisms or specific mammalian tissues.

One field in which the two-hybrid system has been applied with considerable success has been the study of oncogenes and tumor suppressors and the related area of cell cycle control. For example, reconstruction experiments with previously cloned proteins indicated that interactions occur between Ras and the protein kinase Raf (220, 249), human Sos1 guanine nucleotide exchanger and the growth factor receptor-associ-

BEST AVAILABLE COPY

ated protein Grb2 (30), and Raf and the transcription factor inhibitor IκB (129). Two-hybrid searches with oncoproteins or tumor suppressors as targets have identified a leucine zipper protein that binds to Jun (34); protein phosphatase PP1α2, which binds to Rb (55); a bHLH-zip protein Mxi1, which binds to the Myc-associated protein Max (247); and the Rb-related protein p130, which binds to cyclins and was identified through its interaction with the cyclin-dependent kinase Cdk2 (83). A notable convergence of different approaches came about with the identification of another protein that binds to Cdk2, a 21-kDa protein termed Cip1, which inhibits the kinase activity (85). This protein turned out to be identical to a protein encoded by the major p53-inducible transcript (58), suggesting that the tumor suppressor role of p53 may be mediated by its activation of the gene for this 21-kDa protein.

The two-hybrid system has several features that make it useful for analysis of protein-protein interactions. It is highly sensitive, detecting interactions that are not detected by other methods (see, e.g., references 127 and 220). On the basis of binding of different proteins to the retinoblastoma protein, Durfee et al. (56) estimate that the minimal binding constant required to detect an interaction in their version of the two-hybrid system is on the order of 1 μM. This value suggests that the system should be applicable to a wide range of protein interactions. However, it is clear that the minimal affinity interaction detectable will depend on such variables as the level of expression of the hybrid proteins; the number, sequence, and arrangement of the DNA-binding sites in the reporter gene(s); and the amount of reporter protein required for a detectable phenotype. Given these variables, it is likely that some versions of the system may detect weak interactions with binding constants considerably greater than 1 μM. Another advantage is that the interactions are detected within the native environment of the cell and hence that no biochemical purification is required. The use of genetic-based organisms like yeast cells as the hosts for studying interactions allows both a direct selection for interacting proteins and the screening of a large number of variants to detect those that might interact either more or less strongly. With a reporter gene such as the yeast *HIS3* gene, the competitive inhibitor 3-aminotriazole can be used to directly select for constructs which yield increased affinity.

The two-hybrid system is limited to proteins that can be localized to the nucleus, which may prevent its use with certain extracellular proteins. Proteins must be able to fold and exist stably in yeast cells and to retain activity as fusion proteins. The use of protein fusions also means that the site of interaction may be occluded by one of the transcription factor domains. Interactions dependent on a posttranslational modification that does not occur in yeast cells will not be detected. Many proteins, including those not normally involved in transcription, will activate transcription when fused to a DNA-binding domain (134), and this activation prevents a library screen from being performed. However, it is often possible to delete a small region of a protein that activates transcription and hence to remove the activation function while retaining other properties of the protein.

Other Library-Based Methods

A number of other library strategies have been developed recently that, in principle, should result in th identification of proteins that interact with a protein of interest. However, because the first description of methods generally involves known combinations of proteins, th general applicability of a new method cannot be easily judged.

In one approach, the ability of th *E. coli* bacteriophage λ repressor to dimerize was used as a reporter for the int raction of leucin zipper domains (98). Th N-terminal domain of repressor binds to DNA but dimerizes inefficiently; a separate C-terminal domain that mediates dimerization is required for efficient binding of the protein to its operator. The N-terminal DNA-binding domain was fused to the leucine zipper of the yeast Gcn4 protein, which allowed dimerization and repression of transcription in *E. coli*. This repression enabled the host cell to survive superinfection by λ phage. This phenomenon enabled Hu et al. (98) to introduce single-amino-acid mutations into the leucine zipper domain and to use a genetic assay in *E. coli* to determine whether dimerization of the zipper domain occurred. They suggested that this assay could be used to select clones from a library for proteins that bind to a target protein, which is expressed in *E. coli* as a repressor hybrid. Any phage that express a protein that binds to the target protein should compete for dimerization of the repressor and its ability to bind λ operators. These phage would be detected because they result in plaques. As described, this approach would be limited to target proteins that homodimerize. In addition, this method when applied to library screening is a competition assay; it would require that the library-encoded protein bind to the target protein in preference to the target protein interacting with itself.

Another *E. coli*-based assay involves tagging the target protein with biotin by fusing it to the biotin carboxylase carrier protein (74). This tag allows the protein to be bound by avidin, streptavidin, or anti-biotin antibody-coated filters. Potential interacting proteins are fused to the LacZ protein and expressed from a λ vector such that β-galactosidase activity is intact. These phage are infected into cells containing the biotin-tagged target protein, and interaction can occur in vivo between a library-encoded protein and the target protein. This interaction is then detected when the phage plaques are transferred to avidin filters and assayed for β-galactosidase activity. The method was shown to work by using biotinylated c-Jun protein and a c-Fos–LacZ fusion. Although the protein-protein interaction occurs within the living bacterial cells, th detection of this interaction occurs in vitro on filters that must be washed after transfer of the proteins. Thus, in principle, this method may have many of the same limitations that protein probing of expression libraries has.

GENETIC METHODS

For organisms for which powerful genetic analysis methods exist, sophisticated strategies can be designed to uncover genes that show interactions with other genes. In many cases, these newly uncovered genes encode proteins that physically interact with proteins encoded by the known genes. In oth r cases, genetic methods can be used to confirm interactions among previously identified proteins. These strategies are generally based on classical genetic approaches. For example, identification of extragenic suppressors often reveals mutations in genes whose products physically interact with the protein containing the original defect. Synthetic lethal screens yield mutations that, in combination with another nonlethal mutation, result in the inability of the organism to grow; this phenotype is commonly due to alterations in interacting prot ins. Overproduction of certain proteins can lead to th suppression of mutations in interacting proteins. In other cases, overproduction disrupts a cellular process by altering th balance of the different components of a complex structure, or th overproduced protein is nonfunctional and acts in a dominant-negative manner

BEST AVAILABLE COPY

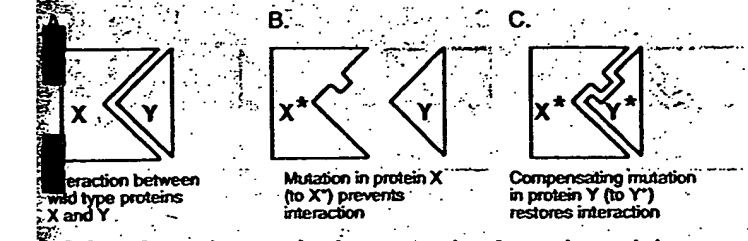


FIG. 7. Extragenic suppression due to restoration of a protein-protein interaction.

The value of some of these genetic approaches has been nificantly increased by applying them to organisms not ame- able to classical genetic techniques, using modern molecular tools. For example, the ability to generate mice either carrying vel genetic information or deleted for one or more of their dogenous genes allows this organism to be analyzed by some of the logic formerly reserved for much simpler creatures. However, it must be kept in mind with any genetic approach at identification of mutants with the correct phenotypes does t guarantee that the biochemical mechanisms invoked to explain these phenotypes are correct.

Extragenic Suppressors

Suppressor mutations are mutations that partially or fully revert the phenotype caused by an original mutation (see reference 86 for review). Extragenic suppressors occur in genes her than the gene carrying the primary mutation. This is strated in Fig. 7, in which a mutation of protein Y to Y* compensates for the defect X* to restore activity to the XY dimer. However, analysis of these suppressors is often difficult, cause they lack any phenotype in the absence of the primary tation. To circumvent this problem, Jarvik and Botstein (107) sought suppressors of temperature-sensitive mutants of hage P22 that resulted in a cold-sensitive phenotype. This ld-sensitive phenotype did not necessarily depend upon the esence of the original mutation causing temperature sensitivity, and thus mutations in new genes could be uncovered. It was proposed (107) that one mechanism of this suppression is at the original mutation and the suppressor lie in genes hose products physically interact and that the original mutation destroyed this interaction. The suppressor then produces a compensating alteration that restores the interaction.

This type of suppressor analysis has been exploited in studying fundamental processes in yeast cells, particularly cell cycle control, cytoskeleton structure, and RNA splicing. Moir et al. (152) isolated cold-sensitive cell division cycle (*cdc*) mutants of *ccharomyces cerevisiae* and used them to identify tempera- re-sensitive revertants. Some of these revertants carried new mutations that alone resulted in a *cdc* phenotype at the restrictive temperature, suggesting that the mutated gene products ght interact with the cold-sensitive protein. These results pport th idea that only a few genes might be capable of mutation to generate an altered product that can suppress the original mutation. Thus, this approach can be applied to a ocess such as cell cycle control and reveal most or all of the teracting gene products.

In a similar strategy, suppressors of a temperature-sensitive tation in the *S. cerevisiae* actin gene that acquired a cold- nsitive phenotype identified five new genes (160). Mutations these genes, even in a background with the wild-type actin gene, led to phenotypes similar to those of actin mutants. These results suggested that these genes could encod proteins at ar part of th actin cytoskeleton. In a related approach, dominant suppressors of an actin mutation also identified a gene whose product may interact with actin (3). In both these cases, th suppressor mutations showed allel specificity; some but not all actin alleles were suppressed by a given mutation. This allele specificity also supported the idea of a direct physical interaction, in that suppressor mutations that simply bypass the requirement for the protein containing the original mutation would not be expected to show such specificity.

The nematode *Caenorhabditis elegans* has also been used extensively for suppression analysis because large populations of individuals can be examined (96). If a temperature-sensitive mutant is available, it can be shifted to the restrictive temperature to apply a direct selection for suppressors. This approach has been used to study such processes as movement, egg laying, and sex determination. One example is the suppression of an *unc-22* mutation that resulted in muscle twitching (151). Some of these suppressors were mutations in the *unc-54* gene which encodes the major myosin gene. These results suggested that the *unc-22* and *unc-54* proteins physically interact, and this idea is supported by the finding that the *unc-22* protein, like myosin, is located in the A-bands of muscle (150).

Suppressor analysis can clearly uncover new mutations that affect a process under study, and analysis of the genes and proteins defined by these mutations sometimes indicates interacting proteins. While often used with temperature-sensitive and cold-sensitive mutations, many other types of spontaneous mutations can also be readily suppressed if an appropriate genetic selection is available. With the availability of numerous cloned genes, conditional alleles can now be generated by in vitro mutagenesis methods. An obvious limitation of this type of analysis is that it can generally be applied only to simple organisms such as phages, bacteria, yeasts, nematodes, and *Drosophila* species. It requires not only the gene of interest but also a useful mutant to initiate the analysis. For example, suppressors in an interacting protein may be difficult or impossible to obtain if the original mutation does not affect a domain of interaction. Furthermore, other mechanisms can yield suppressors. These include second intragenic mutations, gene duplication of the original mutant gene, suppression by epistasis, and informational suppression (see, for example, reference 96). Thus, identification of the suppressors of interest against a background of these other mutations can be a time-consuming process.

Synthetic Lethal Effects

Mutations in two genes can cause death (or another observable defect) while mutation in either alone does not. This phenomenon is called a synthetic effect and can result from physical interactions between two proteins required for the same essential function. This is illustrated in Fig. 8, in which the dimer XY is required for some function and loss of this function results in a detectable phenotype. Mutation in X or Y yields partial binding, but the double mutant X*Y* has no binding. Dobzhansky (52) first described synthetic lethal effects in *Drosophila* species. However, the search for synthetic lethal effects has been applied successfully most often in *S. cerevisiae*. One of the tools available for research in this organism is a colony-sectoring assay (93, 119), in which cells containing a plasmid are red and can therefore be easily distinguished from those that hav lost the plasmid and are white. If maintenance of the plasmid is not essential for viability of the yeast, colonies appear with red-and-whit sectoring. If the cells become dependent on a gene carried by the plasmid, th colonies appear uniformly red. For example, Bender and Pringle (15) used such an assay with a plasmid-borne copy of the *MSB1* gene, which

BEST AVAILABLE COPY

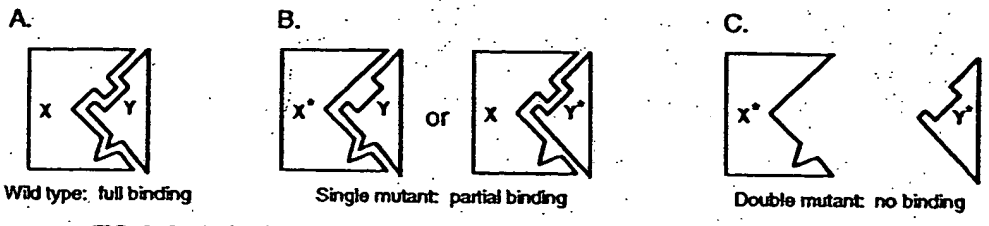


FIG. 8. Synthetic effect, in which either single mutant is functional but not the double mutant.

plays a role in bud formation. They mutagenized the plasmid-containing cells and screened for mutants in which *MSB1* had become essential for survival. This screen identified two new genes, *BEM1* and *BEM2*, in which mutations led to defects in cell polarity and bud emergence. In this approach, if the plasmid is maintained at high copy number in *S. cerevisiae*, it is also possible to identify mutations in new genes that are lethal but can be suppressed by multiple copies of the plasmid-borne gene.

A similar approach was taken by Costigan et al. (40) to identify mutants that require the Spa2 protein, which is also involved in polarized cell growth as well as in the morphogenetic changes that occur in yeast mating. The synthetic lethal screen identified the *SLK1* gene, which is necessary for morphogenesis in vegetatively growing yeast cells and in mating pheromone-treated cells. Costigan et al. pointed out that the synthetic lethal screen by the colony color assay is extremely sensitive and can identify mutants with low viability. Since both *spa2* and *slk1* mutants are individually healthy, the screen did not simply combine two mutations each causing unhealthiness to result in death, a common concern in using this method. Instead, it seems likely that the synthetic lethal effect often results from two different defects in the same cellular process.

Other synthetic lethal screens in yeast cells involve a poison assay in which the presence of a plasmid-borne gene on a particular medium is lethal; when yeast cells containing this plasmid are placed on such a medium, there is strong selection for cells that have lost the plasmid. However, mutants that cannot survive without the plasmid can be identified, because the plasmid also contains the gene of interest whose presence is required in these mutants. Such mutants do not grow on replica plates containing the poison. This approach was used to identify mutations in the 3-hydroxy-3-methylglutaryl coenzyme A reductase genes (12). Alternatively, the gene of interest can be expressed by using a regulated promoter, such that mutants that do not survive the repressed condition are identified. Inducible expression of the yeast *RAS2* gene led to the identification of mutations in the *CYR1* gene, which encodes adenylate cyclase (149). Finally, synthetic lethal effects can be uncovered by combining mutations identified in other genetic screens. For example, yeast cells containing a temperature-sensitive mutation in the *SEC4* gene, essential for secretion, are inviable at the permissive temperature when they also contain a temperature-sensitive mutation in certain other *SEC* genes (190). Yeast cells with mutations in both α-tubulin and β-tubulin are inviable (101).

While synthetic lethal screens often lead to the identification of interacting gene products, other explanations do not require this physical interaction (101). For example, the two proteins might both be components of the same structure, or one protein could regulate the activity of the other. Additionally, there are likely to be some cases in which the combination of two mutations, either of which causes poor growth on its own, leads to complete inviability.

Overproduction Phenotypes

Overproduction of wild-type proteins. The overproduction of some wild-type proteins can lead to phenotypes that provide insight into protein-protein interactions. In *S. cerevisiae*, a multicopy plasmid often suppresses mutations in genes other than the one carried on the plasmid (reviewed in reference 182). For example, a temperature-sensitive mutation in the *CDC28* gene, which encodes a protein kinase involved in controlling cell division, can be suppressed by multicopy plasmids carrying the *CLN1* or *CLN2* gene, which encode cyclins (82).

In other cases, overproduction of a protein can cause a phenotype that is altered by overproduction of an interacting protein. High-copy-number plasmids expressing either of the yeast histone pairs H2A and H2B or H3 and H4 caused an increased frequency of chromosome loss (142). However, overproduction of both pairs of histone proteins did not affect the fidelity of chromosome transmission, indicating that it is the imbalance of the two dimer sets with respect to one another that affects this fidelity (142). Overproduction of the yeast Gal4 protein, the transcriptional activator of the galactose-inducible genes, leads to galactose-independent transcription. However, proper regulation is restored if the Gal80 protein, a negative regulator that binds to the Gal4 protein, is also overproduced (159). While the phenotype due to an overproduced wild-type protein may reflect interactions with another protein (either mutant or wild type), there are several other mechanisms by which such phenotypes can occur. For example, an overproduced protein may bypass the transcriptional regulation due to another protein. In other cases, an overproduced protein may lead indirectly to the stabilization of a mutant protein.

Overproduction of mutant proteins. Overproduction of a nonfunctional version of a protein can result in a mutant phenotype due to disruption of the activity of the wild-type protein (Fig. 9) (reviewed in reference 90). The existence of such dominant-negative proteins can lead to a definition of the oligomerization domain of a protein. An early example of this came from studies of the *E. coli* Lac repressor, which has distinct domains for DNA binding and for oligomerization. A mixed oligomer of wild-type subunits and mutant subunits unable to bind DNA results in a nonfunctional repressor (143). This kind of mutant provides evidence for the multimeric nature of the repressor, and analysis of the sites of mutation defined the domains involved in DNA binding and in oligomerization.

A similar mechanism may operate in many human cancers. The wild-type p53 protein is a transcriptional regulator which is tetrameric, and its oligomerization domain is near the C terminus. Mutations in the central domain of p53 that occur in tumors produce dominant-negative mutant proteins that bind to and inactivate the function of the wild-type protein (67). The ability to manipulate cloned genes and reintroduce these mutant versions into cells now allows dominant-negative mutants to be created in many different organisms. For example,

BEST AVAILABLE COPY

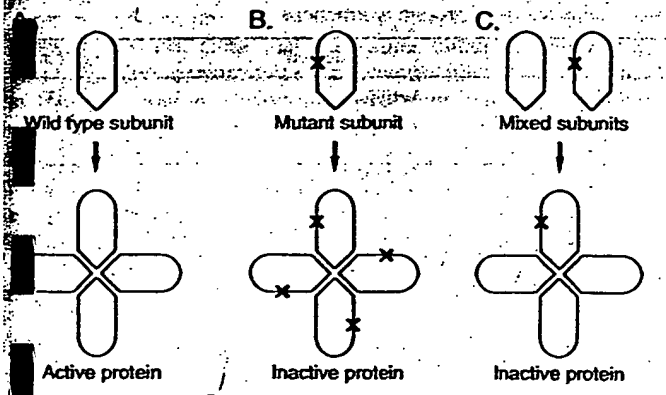


FIG. 9. Dominant-negative effect. Pure populations of wild-type (A) or mutant (B) subunits result in an active or inactive protein, respectively. A mixture of the two types (C) will also be inactive if the mutant subunit acts in a dominant fashion.

dominant-negative Myc proteins were overexpressed in fibroblasts and shown to inhibit transformation by the v-*abl* and *BCR-ABL* oncogenes (194). It was suggested that this effect was due to the mutant Myc proteins competing with the endogenous wild-type Myc protein for binding to the Max protein, thus forming nonfunctional heterodimers.

Unlinked Noncomplementation

Individuals heterozygous for two different recessive mutations sometimes display a mutant phenotype. This unlinked noncomplementation is often interpreted as being due to mutation in two genes that encode interacting products. In *Drosophila* spp., new recessive mutations were identified that failed to complement β_2-tubulin mutations and that mapped to other genes (176). At least one of these mutations mapped very close to an α-tubulin gene. A model for this noncomplementation is based on a minimal dosage requirement for the product of two interacting proteins. If the mutant proteins assemble randomly with the wild type, the double heterozygote would contain only one-fourth the normal level of complex, which would be insufficient for function. In addition, when homozygous, some of the second-site noncomplementing mutations lead to defects in tubulin function, and this property is consistent with the model.

POPULAR METHODS TO ESTIMATE AND DETERMINE BINDING CONSTANTS

Importance of Characterization of the Binding Interaction

The ultimate goal of studying protein-protein interactions is to understand the consequences of the interaction for cell function. This depends in turn on understanding the strength of the interaction in the cell. The determination that two proteins can interact with one another is only the first step in understanding if, and to what extent, the interaction takes place in vivo. Evaluation of the interaction requires the assessment of at least six parameters, which are discussed below.

Binding constant. For any simple interaction of one protein (P) with another (L, for ligand), the interaction is governed by the binding constant K_d, according to the simple equation $K_d = [P_f][L_f]/[PL]$. In this equation, $[P_f]$ and $[L_f]$ refer to the free (i.e., unbound) concentrations of P and L respectively. The interaction between protein and ligand is also expressed in two other ways. First, it is often expressed instead as an affinity constant, $K_a = [PL]/[P_f][L_f]$, i.e., $K_a = 1/K_d$. Second, it is often expressed as a ratio of two rate constants. The rate of formation of PL is $k_a [P_f][L_f]$, where k_a is the association rate constant, and the rate of breakdown of PL is $k_d [PL]$, where k_d is the dissociation rate constant. At equilibrium, the rate of formation of PL equals the rate of breakdown of PL, and $K_d = k_d/k_a$. Evaluation of the dissociation constant is the subject of this section.

Concentrations of species. To evaluate the extent to which two proteins can interact, the cellular (or compartmental) concentrations of P_t (the sum of bound and unbound concentrations) and L_t are required, in addition to the dissociation constant. These two parameters can drastically alter an evaluation of the population of molecules in a complex. For example, if $K_d = [P_t] = [L_t]$, 38% of the species are in the complex PL at any one time. If K_d is 10-fold higher (weaker binding), only 8.4% of the species are in the complex at one time, and if K_d is 10-fold lower (stronger binding), 73% of the species are in a complex. A similar effect holds for alterations in the concentrations of P and L in the cell. A simple way of calculating [PL] from the easily measured parameters $[P_t]$ and $[L_t]$ is as follows: $[PL] = \{([P_t] + [L_t] + K_d)/2\} - 1/2 \{([P_t] + [L_t] + K_d)^2 - 4[L_t][P_t]\}^{1/2}$ (54).

Influence of competing proteins. Even if a protein has high affinity for a ligand protein, L, and the protein and ligand are present in sufficient quantities to interact functionally in the cell, they may not do so in vivo to the same extent as in vitro. Other ligands may effectively compete for the ligand protein if they are present at high enough concentration and interact with sufficient affinity. For example, if the concentration of P and L1 are both equal to the dissociation constant, 38% of the species are in a complex. If another ligand, L2 (or a set of potential ligands), is present at 1,000 times the concentration of L1 and has 10-fold-lower affinity for P, the interaction of P with L2 will titrate the vast majority of the protein P (99%, if L2 was the only interacting protein), leaving very little to interact with L1. This sort of consideration is addressed in part by protein affinity columns, coimmunoprecipitation experiments, and cross-linking, since all the proteins in the applied extract have equal opportunity to bind. It is not addressed in affinity blotting or library-based detection methods, in which gene products are tested individually.

Influence of cofactors. Two types of cofactors can influence protein-protein interactions. First, small effector molecules and ions such as ATP, GTP, and Ca^{2+} can influence many protein-protein interactions. Second, other macromolecules (DNA, RNA, and proteins) can affect protein-protein interactions by forming ternary (or larger) complexes. Such complexes can be very much more stable than the corresponding binary complexes.

Effect of cellular compartmentation. A protein that is interacting with a ligand or a set of ligands is also influenced by its location in the cell. For example, some transcription factors are regulated in part by their partitioning between the cytoplasm and nucleus; they can interact with the transcription machinery only when they are in the nucleus.

Solution conditions. Other factors that can affect the strength of protein-protein interactions include solution conditions (salt concentration, pH, etc.), as well as the effects of molecules such as polyethylene glycol, which causes macromolecular crowding and can significantly lower the observed binding constant of proteins (see, for example, reference 108).

Limits of Binding-Constant Considerations

The lower limit for the concentration of a protein in an organism of the size of the yeast *S. cerevisiae* is 0.1 nM (as-

BEST AVAILABLE COPY

TABLE 1. Dissociation constants for some well-defined protein-protein interactions

Complex[a]	K_d (M)	Method[b]	Reference(s)
PDEαβ:PDEγ	1.3×10^{-10}	Activity	103
	5×10^{-11}	fl. an.	25
	1×10^{-11}	fl. an.	234
	$<1 \times 10^{-11}$	Activity	233
TαGTPγS:PDEγ	$<1 \times 10^{-10}$	int. fl.	164
TαGDP:PDEγ	3×10^{-9}	int. fl.	164
CAP cAMP:RNA polh	3×10^{-5}	fl. an.	91
	1×10^{-6}	fl. an.	170
T7 gene 2.5 protein:T7 DNA polymerase	1.1×10^{-6}	fl. an.	115
λ repressor (dimer to tetramer)	2.3×10^{-6}	fl. an.	9
λ repressor (monomer:dimer)	2×10^{-8}	l.z. gf.	13, 193
Citrate synthase: malate dehydrogenase	1×10^{-6}	fl. an.	214
C4 binding protein: human protein S	6×10^{-10}	Solid phase	158
p85 (PI3K): tyrosine-phosphorylated peptide from PDGF	5.2×10^{-8}	SPR	166
CheY:CheA	3×10^{-8}	SPR	197
CheA:CheW	1.3×10^{-5}	eq. gf.	72
VAMP2:syntaxin A	4.7×10^{-6}	SPR	27
EGF:EGF receptor	4.1×10^{-7}	SPR	249
PKA-C:PKA-R	2.3×10^{-10}	SPR	88
PRI:angiogenin	7×10^{-16}	Fluorescence, exch	126
ras:raf	5×10^{-8}	GST ppt'n	227
NusB:S10	1×10^{-7}	Sucrose gradient sed'n	138
NusA: core RNA polymerase	1×10^{-7}	Sucrose gradient sed'n	80
		Fluorescence tag	76
Trypsin:pancreatic trypsin inhibitor	6×10^{-14}	Kinetics, comp'n	221

[a] Abbreviations: PDE, phosphodiesterase; TαGTPγS, α subunit of transducin complexed with GTPγS; TαGDP, α subunit of transducin complexed with GDP; CAP cAMP, catabolite gene activator protein complexed with cAMP; RNA polh, RNA polymerase holoenzyme; PDGF, platelet-derived growth factor; VAMP2, vesicle-associated membrane protein 2; PKA-C, catalytic subunit of protein kinase A; PKA-R, regulatory subunit of protein kinase A.

[b] Abbreviations: fl. an., fluorescence anisotropy; int. fl., intrinsic fluorescence; l.z. gf, large zone equilibrium gel filtration; eq. gf., equilibrium gel filtration; SPR, surface plasmon resonance; exch, exchange; ppt'n, precipitation; sed'n, sedimentation; comp'n, competition.

suming a radius of 1.5 μm and one molecule per cell), and for an animal cell with a radius of 10 μm, the lower limit is about 0.3 pM. Thus, for two such proteins to interact a significant percentage of the time, the dissociation constant must be at the same concentration (in which case they will interact 38% of the time). At the other extreme, some glycolytic proteins represent 1% or more of the soluble protein in the cell. Indeed, glyceraldehyde-3-phosphate dehydrogenase is reported to approach 20% of the soluble protein in *S. cerevisiae* under certain conditions. This upper limit corresponds to 1.7×10^7 protein molecules per cell and a cellular concentration of 1 mM, and it represents the upper limit for binding-constant considerations of two such proteins. In considering protein concentrations, it is worth noting that a typical yeast cell contains about 3×10^5 ribosomes (226), 100 to 500 molecules of tRNA splicing enzymes (169, 178), and 300,000 molecules of actin (157).

Methods for Determining Binding Constants

A number of methods have been described to measure binding constants. Some of the more commonly used ones are described below, together with a brief evaluation of the method. The values of dissociation constants for several protein-protein interactions are listed in Table 1.

Binding to immobilized proteins. Protein affinity chromatography can be used to estimate the binding constant. This method is well described in an excellent review (69). The form of the binding equation that is used in this sort of experiment expresses the fraction of L bound to protein P as follows: $[PL]/[L_t] = [P_f]/([P_f] + K_d)$. As long as th concentration of covalently bound protein $[P_t]$ is in great excess over that of the ligand, $[P_t] \sim [P_f]$ and the fraction of protein L that is bound is $[P_f]/(K_d + [P_f])$. Thus, if $[P_t] = 100\ K_d$, essentially all of L is bound (a little more than 99%), and if $[P_t] = 0.01\ K_d$, very little of L is bound (a little less than 1%).

Columns are prepared with different concentrations of covalently bound protein. Then a preparation of the interacting protein ligand is loaded on the column and washed with 10 column volumes of buffer, and bound protein is eluted with SDS. At a concentration of 20 K_d, the covalently bound protein retains 95% of the ligand in one column volume and therefore 0.95^{10} or 61% in 10 column volumes. Thus, the lowest concentration of bound protein that allows retention of most of the ligand is 20 K_d.

The percentage of bound ligand drops very quickly as the concentration of covalently bound P on the column is lowered, particularly as the concentration of P_t approaches K_d. At 5 K_d 16% of the ligand would be retained, at 2 K_d 1.7% of the protein would be retained, and at 1 K_d only 0.1% would be retained. It is for this reason that detection of interacting proteins by affinity chromatography depends critically on th concentration rather than the amount of bound protein (see the section on protein affinity chromatography, abov).

An important parameter in this experiment is the amount of protein that is active on the column. Estimates range from 10% for gene 32 protein to about 50% for others (69). A second factor is the amount of pure protein available to be coupled. If protein is limiting, sufficiently high concentrations of bound protein on the gel are achieved only with appropriate microcolumns. Such columns, with as little as 20 μl of appropriate beads, are described in detail by Formosa et al. (69). With the recent widespread use of gene fusion technology, larg quantities of protein are not a serious probl m with most cloned structural genes. A third factor, which is evident from the discussion above, is the form of the protein that is used for the determination. Proteins that require modification to be active must be purified in that form for proper evaluation.

BEST AVAILABLE COPY

This method works well in estimating the binding constant. However, it is not clear that the values obtained represent a true equilibrium constant; if so, one would have to assume that the bound ligand is always in equilibrium with the solution ligand during flow of the column and that interactions of solid-phase bound protein with liquid-phase ligand are the same as interactions in the liquid state. Nonetheless, for interactions that have been measured by more than one method, the results agree well (see reference 69 and references therein).

Sedimentation through gradients. The method of sedimentation through gradients measures populations of complexes by monitoring the rate of sedimentation of a mixture of proteins through gradients of glycerol or sucrose. Fractions are assayed by appropriate methods (activity, immunoblotting, etc.) to determine the elution positions of each protein. Proteins will sediment as a complex at concentrations above the binding constant (provided that the complex is stable; see the discussion below) and at their native positions at concentrations below the binding constant. By varying the concentration of one or both of the proteins and taking into account the dilution of the species during sedimentation, one can reasonably accurately bracket the binding constant. For example, the binding constant of *E. coli* NusB protein and ribosomal protein S10 was estimated at 10^{-7} M based on the observation that S10 protein sedimented faster (with NusB protein) when both were at 6×10^{-7} M, slightly more slowly when both were at 3×10^{-7} M, and much more slowly (midway between its sedimentation position alone and its fully complexed sedimentation position) when both were at 1.5×10^{-7} M (138). There are two reasons that S10 sedimented at an intermediate position rather than at its own position during the run at 1.5×10^{-7} M of each protein. First, the proteins are usually about fivefold more dilute at the end of the sedimentation than when they are first loaded on the gradient; therefore, if S10 protein could bind at the beginning of the run (and sediment faster), it might not bind at the more dilute concentration at the end of the run. Thus, it would sediment at an intermediate position. Second, equilibrium binding is a dynamic process and molecules are constantly associating and dissociating. Therefore, an individual S10 molecule which dissociated from NusB at the trailing edge of the peak would be in a region with very much less NusB to bind. It would sediment at its native rate from that point on.

There are two problems associated with this technique. First, it is not an equilibrium determination, because of the changing conditions during the run. Therefore, failure to detect an interaction may be due to rapid equilibrium rather than a lack of interaction. As such, values obtained from this type of experiment represent an upper bound for the binding constant. Second, sedimentation through gradients does not resolve species that well. Sedimentation rates vary as $M^{2/3}$ for spherical molecules. Thus, dimerization of one spherical molecule with one that is 1/10 the mass will increase its sedimentation rate by only 6%, which is very difficult to detect; in contrast, the change in mobility of the smaller molecule will be fivefold under such conditions.

Although this method has limitations, it has been useful for estimating the upper limit of a binding interaction.

Gel filtration columns. Gel filtration is another simple way of estimating the binding constant. In gel filtration, the elution position of a protein or of a protein complex depends on its Stokes radius. This provides a very powerful and conceptually simple method for evaluating the strength of the interaction between two different proteins. Such sizing columns have been used in three distinct ways to measure or to estimate the binding constant.

(i) Nonequilibrium "small-zone" gel filtration columns. In the simplest approach, a solution containing a protein and a ligand protein is applied in a small volume to the column and the material is resolved in the usual way. This is called a "small-zone" column. The elution positions of the protein and ligand in the mixture are compared with those of the protein and ligand when each is chromatographed individually on the same column. If a complex has formed between the protein and ligand, the complex will elute earlier than either protein alone. From measurements of the concentrations of species required to form a complex, one can estimate the binding constant. This type of experiment has been used, for example, to measure the binding of *E. coli* NusA protein to core RNA polymerase and has yielded values very similar to those determined by fluorescence measurements (76). Similarly, Herberg and Taylor (89) quantitated the interaction of cAMP-dependent protein kinase with both the R1 subunit and PKI in the presence and absence of MgATP.

This direct-application method is not an equilibrium method. Since the concentrations of species change during gel filtration (by diffusion and by dilution), the results are subject to the same sources of error as those of sedimentation through sucrose gradients (see references 2 and 250 for a discussion). Thus, the binding constants calculated in this way can be vastly underestimated, particularly if the complex is in rapid equilibrium (see Fig. 3 of Gegner and Dahlquist [72]) for a vivid contrast between nonequilibrium and equilibrium gel filtration). However, several modeling systems have been described (see reference 211 and references therein).

(ii) Hummel-Dreyer method of equilibrium gel filtration. Gel filtration can also be used as an equilibrium method to establish the binding constant between a protein and its ligand protein. One such method is based on the classic paper by Hummel and Dreyer (102). In this gel filtration method, both the gel filtration buffer and the sample had ligand at the same concentration, but only the sample contained protein. Elution of a protein through such a column caused an increase in the concentration of ligand where the protein eluted, followed by a trough of ligand concentration representing ligand that had been removed in the binding. Evaluation of the binding constant of the protein-ligand complex was simply a matter of knowing the concentration of protein eluted, the free concentration of ligand (set by the column), and the concentration of ligand bound with protein (the concentration of ligand in samples containing protein).

This elegant method has been applied to the interaction of two proteins in only a few cases. As illustrated in Fig. 10, the gel filtration buffer contains protein ligand, and the applied sample contains gel filtration buffer (with the same concentration of protein ligand) as well as the other protein. Gegner and Dahlquist (72) used a column equilibrated with CheW to demonstrate and quantitate the interaction of CheA with CheW. They varied the CheW concentration in the initial sample (while maintaining a constant concentration of CheA in the sample and CheW in the buffer) and quantitated the peak area at the CheW position. The CheW concentration in the sample at which there was no resulting CheW peak or trough represented a sample at true equilibrium. From this, they could calculate a dissociation constant of the complex of 13 μM. A similar series of experiments was done by Yong et al. (243) to demonstrate an interaction between glycerol-3-phosphate dehydrogenase and lactate dehydrogenase over an extremely limited range of NADH concentrations. Such a complex was observed only when the NADH concentration was high enough for an interaction and low enough to be shared by the two enzymes, and it provided evidence for substrate channeling.

BEST AVAILABLE COPY

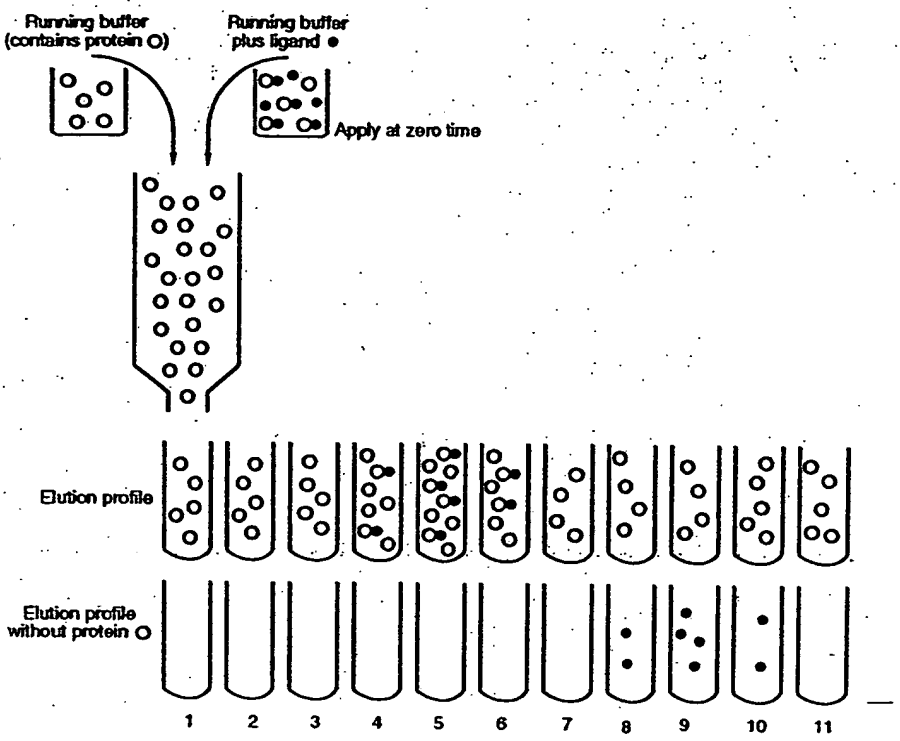


FIG. 10. Equilibrium gel filtration. A solution containing both protein ligand (solid circles) and interacting protein (open circles) is applied to a gel filtration column which is equilibrated with solution containing the interacting protein and developed with running buffer containing the interacting protein. The elution pattern is shown in the first row of test tubes at the bottom. The second row of test tubes indicates the elution pattern that would be observed in the absence of interacting protein.

This method is so simple and inexpensive that it is likely to become much more widely used than at present. Moreover, as an equilibrium experiment, it is without any flaw. The only requirements of this technique are that sufficient protein is available for the experiments and that the elution position of the complex differs from that of at least one of the interacting proteins. With the development of rapid techniques for large-scale protein purification through the use of fusion proteins, it should become relatively routine to obtain enough of any protein to use as a column eluant.

Another variation of Hummel-Dreyer columns is the partitioning method. In this technique, a protein and its ligand protein are mixed with a gel and allowed to equilibrate and the gel is centrifuged or filtered to separate the aqueous phase. From an analysis of the distribution of the protein and the ligand protein in the filtrate and in the gel when they are added separately or together, the K_d can be calculated. An example of this technique is the demonstration of a complex between transaminase and glutamate dehydrogenase which occurs with a dissociation constant of 16 to 67 μM, depending on the presence of various metabolites (63); this is another example of metabolite channeling. This method is also not in wide use, although it seems simple and accurate.

(iii) Large-zone equilibrium gel filtration. One final method of equilibrium gel filtration is the large-zone method (1, 2), in which a very large sample volume is applied to the column, followed by conventional buffer elution. Because a large volume is applied, the concentration of the eluted protein is fixed and constant during th experiment, except at the leading and trailing edges. Th elution position of the leading or trailing edge (which measures th size of th molecule) is then monitored as a function of the sample concentration applied to the column. From such experiments, calculation of th dissociation constant is thermodynamically rigorous, as it is for the Hummel-Dreyer method. This large-zone method has been used to monitor self-association of proteins as well as interactions of dissimilar subunits (see, for example, references 75 and 122), but it has received only limited attention because of the larg amounts of protein needed to do the experiments.

A variation of this method, first described by Sauer (193), monitors the change in elution position of radiolabeled protein mixed with different concentrations of unlabeled protein in different runs. The use of labeled protein allows simpler and more accurate determination of the elution position, thus allowing Sauer to determine a dimerization constant of 20 nM for repressor. Improvements in protein labeling have demonstrated that the lower limit of detection for this method is a K_d of the order of 10^{-12} M (13).

Sedimentation equilibrium. Although sedimentation equilibrium is a classical method of determining the molecular weight of a protein, it has not been widely used to study protein-protein interactions. However, recent progress makes this method much more accessible on a day-to-day basis (see reference 185 for a recent review). Sedimentation equilibrium can now be done in everyday preparative ultracentrifuges with swinging-bucket rotors, and samples can be readily collected because of the development of a highly reproducible BRANDEL microfraction collector (183). These developments allow the use of a variety of techniques to assay the protein content of each sample, including kinetic assays, radioactive tracers (183), and gel analysis of samples (47); the result is a huge increase in sensitivity over that obtained with the old model E centrifuge (184).

Fluorescence methods. Since fluorescence is a highly sensitive method for detecting prot ins through their tryptophan residues, it is potentially a useful way of evaluating protein-protein interactions. Two such m thods have been used and are described below.

BEST AVAILABLE COPY

(i) Fluorescence spectrum. Changes in the fluorescence ission spectrum on complex formation can occur either by a ft in the wavelength of maximum fluorescence emission or by a shift in fluorescence intensity caused by the mixing of two proteins. Therefore, the fluorescence intensity at a particular velength can be used to evaluate the dissociation constant. good example of this technique is illustrated by th interaction of the γ subunit of cGMP phosphodiesterase (PDEγ) subunit with the transducin α subunit (Tα) in the presence of TPγS or GDP (164).

An equimolar solution of TαGTPγS and PDEγ causes a blue shift in the fluorescence emission spectrum relative to the sum of the individual fluorescence spectra, resulting in a difference spectrum [F (complex) − F (sum)] with a positive component at low wavelengths (320 nm) and a negative component at higher wavelengths (357 nm). Titration of PDEγ into a solution of TαGTPγS therefore caused an enhanced increase the fluorescence at 320 nm relative to that observed by ration of PDEγ into buffer alone (and a corresponding decrease at 357 nm) until the TαGTPγS was all complexed, after which further addition PDEγ caused no changes in fluorescence intensity relative to that observed in buffer alone. When rrected for PDEγ fluorescence, both curves yielded the same binding curve, and the K_d for the interaction was evaluated at <100 pM. The interaction of TαGDP with PDEγ results in a ge increase (ca. 70%) in the intensity of the fluorescence ission spectrum relative to the sum of the individual spectra, and this was used to evaluate the K_d at 2.75 nM.

This technique has two limitations. First, the probability of tecting a change in the fluorescence spectrum decreases with e total number of tryptophan residues in the two proteins, since the fluorescence spectrum is the sum of the contributions from all the tryptophan residues. Since PDEγ has only one ptophan residue and Tα has two, this condition was easily t in studying the Tα-PDEγ complex. Second, the sensitivity is limited by the intensity of the fluorescence change, which in turn depends on the inherent sensitivity of fluorescence (of the der of nanomolar) and the change that is observed (which is t easily predictable). Thus, the binding constant was too low to evaluate the TαGTPγS-PDEγ interaction (<100 pM) but was high enough to evaluate the interaction in the presence of DP (2.75 nM).

Although these two limitations exclude the study of many interactions, a number of proteins have a small or limited number of tryptophan residues. For example, bovine Hsc70 s only two tryptophans, and its interaction with small peptides has been evaluated because of the resulting quenching of the fluorescence intensity (123). Similarly, the interaction of angiogenin (one tryptophan) with human placental RNase inhibitor (six tryptophan residues) causes a 50% increase in orescence (126), and the dissociation of mitochondrial creatine kinase (four tryptophans per monomer) from octamers to dimers results in a 25% decrease in fluorescence (81).

A second way in which fluorescence is used to measure the eraction of proteins is with a fluorescent tag. This allows for greater sensitivity of monitoring interactions, as long as the fluorescent adducts do not adversely affect the function of the odified protein or its interaction with other proteins. An ample of this approach is the interaction of spinach calmodulin with smooth myosin light-chain kinase (146). Calmodulin from spinach has a single cysteine, which could be quantitively labeled with 2-(4-maleimidoanilino)-naphthalene-6-fonic acid (MIANS). Calmodulin labeled with MIANS was as efficient as the wild type in activating calcineurin, in activating cGMP-dependent phosphodiesterase, and in binding ter- m. The fluorescence of MIANS-labeled calmodulin increased 80% on binding calcineurin, more than fourfold when bound with myosin light-chain kinase, and twofold on binding caldesmon. In each case, the fluorescence chang required the presence of calcium, and titrations were done to measur the K_d (<5, 9, and 250 nM, respectively).

(ii) Fluorescence polarizati n or anisotropy with tagged m lecules. Because of the long lifetimes of excited fluorescent molecules (nanoseconds), fluorescence can also be used to monitor the rotational motion of molecules, which occurs on this timescale. This is accomplished experimentally by the use of plane-polarized light for excitation, followed by measurement of the emission at parallel and perpendicular planes. Since rotational correlation times depend on the size of the molecule (approximately 1 ns/2,400 Da for an idealized molecule), this method can be used to measure the affinity of two proteins for one another because of the increased rotational correlation time of the complex. Fluorescence anisotropy is done most often with a protein bearing a covalently added fluorescent group, which increases both the observed fluorescence lifetime of the excited state and the intensity of the fluorescent signal.

A good example of this technique is described by Weiel and Hershey (229), who studied the interaction of protein synthesis initiation factor 3 (IF3) with 30S ribosomal subunits by using fluorescein-labeled IF3. The labeled protein routinely had about one dye molecule per monomer, and most of the IF3 protein had one or two dye molecules attached. Fluorescein-labeled IF3 was biologically functional: it bound 30S ribosomal subunits, as measured by sucrose density gradients, at a saturable site(s) and had 80 to 100% of the activity of the native protein in stimulating binding of $tRNA^{Met}$ to 70S ribosomes in the presence of RNA. In the presence of 30S ribosomes, both the fluorescence emission spectrum and the fluorescence lifetime of the fluorescein-labeled IF3 were unchanged. Thus, the observed increase in fluorescence polarization which was associated with binding of 30S ribosomes was most consistent with the expected change in polarization as a result of binding a larger molecule. The Scatchard plot derived from the polarization data gave a stoichiometry of 1:1, and the dissociation constant from the polarization data was 3.2×10^{-8} M. Moreover, wild-type nonderivatized IF3 competed for the binding site with the same binding constant. Thus, the fluorescent probe had no effect on any measurable parameter and the measured K_d is likely to be accurate.

Similar experiments have been done with a variety of systems to evaluate the strength of protein-protein interactions. Fluorescein-labeled IF2 was slightly less active than nonderivatized protein, and the binding to 30S ribosomes was twofold weaker than that of the corresponding unlabeled protein (230). T7 gene 2.5 protein labeled with near-molar amounts of fluorescein isothiocyanate caused both a decrease in fluorescence and an increase in anisotropy when bound with T7 DNA polymerase. The fluorescein isothiocyanate-modified protein had no effect on activity, and the binding constant determined by anisotropy (1 μM) was nearly the same as that determined by anisotropy measurements of EDANS-labeled gene 2.5 protein (1.3 μM), for which the rotational correlation time indicated a 1:1 complex (115). The interaction of (florescein-labeled) citrate synthase and malate dehydrogenase was shown to be well within th physiological range (K_d = 1 μM) and varied as much as 25-fold in th presence of different metabolites (214). The tetramer-dimer equilibrium of λ repressor could be observed with dansylated λ repressor, because of its long fluorescence lifetime and high anisotropic value (indicating rigid orientation), but not with fluorescein, which was attached in the highly

BEST AVAILABLE COPY

mobile N-terminal arm of the repressor molecule (and therefore gave low values) (9).

A variation of this technique has been developed for the interaction of a DNA-binding protein with another protein, in which the DNA is fluorescently labeled (91). In this way, *E. coli* CAP could be shown to interact with RNA polymerase holoenzyme in the presence of cAMP and in th absence of a promoter site. The fluorescently labeled DNA oligonucleotide had a CAP-binding site but no RNA polymerase-binding site, and the resulting increase in polarization allowed the determination of a CAP-RNA polymerase binding constant (2.8×10^{-7} M). Since this interaction was not observed with a CAP mutant protein that was defective in transcription activation, it seems likely that the interaction is important physiologically. Other fluorescent polarization experiments suggest that the CAP-RNA polymerase interaction is much stronger in the presence of cAMP and requires σ factor (170).

Solution equilibrium measured with immobilized binding protein. A simple technique for measuring the dissociation constant of a solution of interacting proteins makes use of bound competitor protein to determine the amount of free protein in such a solution. This method was first described for antibody-antigen reactions (71) and later modified for general use to determine the interaction of C4b-binding protein (C4BP) with human protein S (HPS) (158). A solution containing C4BP and HPS was incubated until equilibrium was reached. The amount of free C4BP in the solution was then determined by incubating an aliquot on a plate containing immobilized HPS under conditions (short incubation time) in which a limited amount of the free C4BP binds the immobilized HPS. This resulted in little perturbation of the equilibrium during the assay for C4BP retained by the immobilized HPS, which was quantitated by an antibody-based method.

This method requires satisfaction of three criteria. First, the two proteins (HPS in solution and HPS immobilized on the plate) cannot bind each other. If they did, C4BP could be captured through HPS-HPS interactions. Second, HPS in solution and HPS immobilized on the plate must compete for the same binding site. This is obviously true in this case, but it is not necessarily true if, for example, anti-C4BP is used in the immobilized system to detect the amount of free C4BP. Third, the method requires that only free C4BP be measured during the incubation with immobilized HPS. This in turn requires that binding to the immobilized HPS remove only a small portion of the total C4BP (<10% was removed in this example) so that equilibrium of the solution is perturbed as little as possible. This condition also requires that the off rate of the complex is low compared with the time of incubation with the immobilized HPS; otherwise, HPS-C4BP complexes could dissociate during the incubation with immobilized HPS and the dissociated C4BP would be measured as free C4BP. Thus, this method, although simple, provides only an upper bound of the dissociation constant.

Surface plasmon resonance. The recent development of a machine to monitor protein-protein and ligand-receptor interactions by using changes in surface plasmon resonance measured in real time spells the beginning of a minor revolution in biology. This method measures complex formation by monitoring changes in the resonance angle of light impinging on a gold surface as a result of changes in the refractive index of the surface up to 300 nm away. A ligand of interest (peptid or protein in this case) is immobilized on a dextran polymer, and a solution of interacting protein is flowed through a cell, one wall of which is composed of this polymer. Protein that interacts with the immobilized ligand is retained on the polymer surface, which alters th resonance angl of impinging light as a result of the change in refractive index brought about by increased amounts of protein near the polymer. Since all proteins have the same refractive index and since there is a linear correlation between resonance angle shift and protein concentration near the surface, this allows one to measure changes in protein concentration at the surface due to protein-protein or protein-peptide binding. Furthermor , this can be done in real time, allowing direct measurement of both the on rate and the off rate of complex formation. A good layman's review of surface plasmon resonance is found in articles by Malmqvist (136) and Jonsson et al. (109), and a clear derivation of the appropriate equations is found in the article by Karlsson et al. (111).

In practice, determination of a binding constant requires measurement of two parameters. First, the increase in RU (resonance units) is measured as a function of time by passing a solution of interacting protein past the immobilized ligand until (usually) the RU values stabilize. Second, the decrease in RU is measured as a function of time with buff r lacking interacting protein. This produces a sensorgram for each concentration of protein, a continuous recording of RU versus time. This procedure is then repeated at a number of protein concentrations, after regeneration of the dextran surface. From these two sets of data, two lines are constructed whose slopes correspond to k_a (the on rate) and k_d (the off rate); from these data, K_d is calculated as k_d/k_a. An alternativ determination of K_d can be made by using the steady-state RU values at different protein concentrations.

This system has several advantages. First, it requires very little material. Typically only 1 to 10 μg of protein has to be immobilized on a sensor chip, which can be reused up to 50 times after removal of adhering protein. Similarly, solutions of interacting protein are in the range of 0.01 to 1 ml, depending on the chosen flow rate (109). Second, the method is very fast. A typical run for a given protein takes about 10 min. Third, no modifications of the proteins are required, such as labeling or fluorescent tags. Fourth, interactions can be observed even in complex mixtures. Fifth, both the on rate and the off rate are readily obtained. Sixth, the system is useful over a wide range of protein concentrations. The practical lower limit of the original Biacore system is a change in resonance angl of 10^{-3} degrees (10 RU), corresponding to surface concentrations of 10 pg/mm^2; moreover, the system is linear up to RU values of 30,000 (109). Seventh, the system is quite sensitive; the practical limit for association rates is 10^6/M/s, and off rates as low as 1.1×10^{-5}/s have been measured by recording for 6 h with buffer (197).

This technique has been used successfully to monitor protein-peptide interactions. A good example is the determination of the binding interaction of different SH2 domains with two tyrosine-phosphorylated substrate peptides derived from platelet-derived growth factor (166). The corresponding peptides were attached to the dextran polymer chip via avidin on the chip and biotin on the peptides. Subsequent real-time analysis demonstrated that interaction of these peptides with the p85 subunit of phosphatidylinositol-3-kinase (PI3K) was characterized by a very high association rate (2×10^6/M/s) and dissociation rate (0.1/s) for the 12-mer peptide Y740P and that most of this binding was contributed by the C-terminal subunit of p85. In this particular case, the dissociation rate of bound p85 had to be determined in the presence of a sink of excess competing peptide in th buffer; otherwise, rebinding of dissociated p85 was a significant problem because of the very high on rate. A similar study of p85 SH2 domain interactions with different tyrosine-phosphorylated peptides (from IRS-1) led to the same conclusions of a high on rate and off rate, which was

BEST AVAILABLE COPY

o measured in th presence of a sink of peptide (64). In this e, the on rate was too high to measure directly (as high as $\times 10^8$/M/s for the C-terminal SH2 domain of p85) and was nstead inferred from steady-state binding and off rat mea- rements and confirmed by competition experiments with e phosphorylated peptide (64). On rates in excess of 10^6/M/s n be limited by mass transport rates (fluid flow through the ell) rather than binding-reaction rates, although this can be rtially compensated by either higher flow rates or a smaller ount of peptide on the chip (111). Competition experiments re also used to show that the affinity of p85 for phosphory- lated peptides was 300- to 800-fold greater than for the corre- onding nonphosphorylated peptide and was as much as 100- d weaker with a glycine or arginine at the +1 position lative to the tyrosine compared with bulky hydrophobic groups or glutamate (64).

One final study demonstrated that a specific threonine res- ue in the SH2 domain of Src, when changed to a tryptophan, ncreased th affinity of the domain for phosphorylated pep- tides which were substrates for GRB2 and that the correspond- g tryptophan of GRB2, when altered to threonine, weakened e affinity of GRB2 for this peptide (137). In each of these ree examples, the primary determinant of specificity was the on rat rather than the off rate.

Surface plasmon resonance has also been used with great ccess to monitor protein-protein interactions. One such ex- mple is the demonstration of a quarternary complex of CheY with CheA, CheW, and Tar (197). CheY was bound to the xtran surface through a unique (and engineered) cysteine sidue, which did not affect chemotaxis activity and which was emote from the interaction domain (197). CheA binds this immobilized CheY protein with a low association rate (368/ /s) and a very low off rate (1.14×10^{-5}/s). Moreover, CheA, eW, and Tar probably form a quaternary complex with eY; addition of all three proteins greatly increases the amount of protein bound to CheY relative to that obtained th CheA alone, although neither Tar nor CheW binds CheY dividually or when present together.

Other examples of protein-protein interactions studied by surface plasmon resonance include the interaction of mono- nal antibodies with human immunodeficiency virus type 1 re protein p24 (111), EGF with the EGF receptor (249), the gulatory and catalytic domains of cAMP-dependent protein kinase (88), and VAMP2 and sytaxin 1A (27).

Two minor problems are associated with surface plasmon sonance measurements. First, immobilization of the ligand otein must be of such a nature that it does not impede or artificially enhance interactions. This is the same problem that associated with protein affinity columns. Attachment of eY was accomplished by using a single site remote from the teraction domain (197); this presents the interacting face to the solvent. Phosphorylated peptides were attached by bioti- lation of the peptide at a single site (but variable position) th a long spacer followed by noncovalent interaction with an vidin-coupled sensor chip (166), and attachment of monoclo- nal antibodies to the chip was accomplished through noncova- nt binding to covalently coupled rabbit anti-mouse IGGFc 11). Primary amines are often linked directly to the dextran polymer, leading to more homogeneous presentation of sur- faces to the solvent but causing possible inhomogeneities in teraction (88). Second, the sensor chip has to be regenerated der conditions which do not denature the immobilized li- gand protein. Protein adhering to the immobilized C subunit of protein kinase A was removed with cAMP (88), proteins bind- g to immobilized phosphorylated peptides were removed th a pulse of dilute SDS (166), and CheY was regenerated with a pulse of guanidin hydrochloride (197). In some cases, the ligand is deliberately removed befor the next experiment; thus, monoclonal antibodies sticking to IGGFc were removed with dilute HCl before readdition of th monoclonal antibod- ies to act as a ligand for p24 binding (111).

Limits to Detection

Determination of the binding constant of tightly interacting species by standard methods described above depends on be- ing able to determine and quantitate the fraction of protein ligand bound at a given protein concentration that spans the dissociation constant. For a standard 50,000-kDa protein, the practical limit of silver staining is of the order of 0.2 ng or 20 μl of a 10-ng/ml solution, which would be useful for a dissoci- ation constant of 1 nM or greater. For in vitro translated protein, the practical limit is 1,000 Ci/mmol times the number of amino acid residues, or 1,000 dpm of ^{35}S-labeled protein per fmol (singly labeled); this corresponds to 10^{-12} M or, with 10 residues incorporated, 10^{-13} M; therefore, allowing for con- centrations below K_d, the lower limit of detection is of the order of 10^{-12} M.

Some protein-protein interactions are too tight ($K_d < 10^{-12}$ M) to measure by the methods described above. For example, human placental RNase inhibitor (PRI) interacts very tightly with both angiogenin ($K_d = 7 \times 10^{-16}$ M) (126, 126a) and human placental RNase ($K_d = 9 \times 10^{-16}$ M) (199). For the interaction of PRI with angiogenin, the association rate con- stant, k_a, was measured by monitoring the change in intrinsic fluorescence by stopped-flow fluorescence techniques, and the dissociation rate constant, k_d, was measured by measuring the release of PRI in the presence of scavenger RNase, to which it binds and inhibits the activity.

A dissociation constant of the magnitude of 7×10^{-16} M for the PRI-angiogenin interactions means that the dissociation rate is measured in weeks! In this case, the $t_{1/2}$ for dissociation of the complex was 60 days (corresponding to $k_d = 1.3 \times 10^{-7}$/s). Furthermore, the overall on rate of 1.8×10^8/M/s liters · mol/s is near the diffusion limit for molecules of the size of proteins. It is hard to imagine what selective pressure would require or maintain such a tight interaction. This is particularly true since human placental RNase and angiogenin both bind PRI equally tightly and are substantially different at the amino acid level.

It is possible that a number of macroscopic protein-protein interactions operate at this level. Any protein composed of three or more subunits can have significant interactions among individual pairs of the component protein. If, for example, a subunit has a K_d of 10^{-7} M with each of two other subunits, the effective K_d of the dissociation of that subunit from the com- plex is 10^{-14} M (see reference 116 for a discussion of this point). Thus, complicated structures like the ribosome might effectively lock the proteins together in undissociable units. It is also possible that other, simpler interactions are this tight; the dissociation rate of the subunits of a number of proteins that purify as a complex tends never to be investigated.

EXAMPLES OF WELL-CHARACTERIZED DOMAINS

Given that a straightforward set of experiments is all that is required nowadays to identify two proteins that interact and to delineate th domains responsible for the binding, toward what ends does this analysis continue? To address this question, it is instructive to consider the case of some domains involved in protein-protein interaction that have been extensively charac- terized. Using a combination of numerous techniques, includ-

BEST AVAILABLE COPY

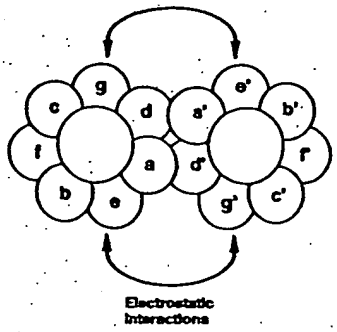


FIG. 11. Helical wheel representation of a leucine zipper. Adapted from reference 221a with permission of the publisher.

ing detailed structural approaches, investigators who have focused on the analysis of leucine zippers, SH2 domains, and SH3 domains have made tremendous advances in the last few years. These studies have considerably extended our understanding of transcriptional regulation and signal transduction. In the next sections, we provide a brief view of how these three domains function.

Leucine Zipper

The leucine zipper is a protein-protein interaction motif in which there is a cyclical occurrence of leucine residues every seventh residue over short stretches of a protein in an α-helix. These leucine residues project into an adjacent leucine zipper repeat by interdigitating into the adjacent helix, forming a stable coiled-coil. This motif was first described by Landschulz et al. (124) in connection with a new structure within DNA-binding proteins that might be responsible for interactions with a similar motif to promote specific DNA binding by basic amino acid residues adjacent to the leucine zipper motif (hence the name bZIP). The leucine zipper model was originally proposed on the basis of the leucine distribution and amino acid sequence of regions of C/EBP, Myc, Fos, Jun, and Gcn4. It is now known to be common to over 30 proteins (59). Subsequent experiments have confirmed the existence of this structure and have extended these observations.

Structure. The X-ray structure of the Gcn4 leucine zipper region (consisting of 33 amino acids) demonstrates that the leucine zipper consists of two parallel coiled coils of α-helices wrapped around each other and forming one-quarter of a turn of a left-handed supercoil (59, 161; also see reference 4). The dimer forms a smoothly bent cylinder about 45 Å (4.5 nm) long and 30 Å (3 nm) wide. On a helical-wheel representation of the α-helix (Fig. 11), the leucines occupy position d (and d' of the adjacent helix) and share the interior with the residues at position a (a'), as well as parts of residues e and g (and e' and g'). The packing corresponds to the "knobs into holes" model proposed by Crick (42), in which each interior amino acid residue is packed into a gap formed by four nearest neighbors from the opposite helix. More than 95% of the surface area that is buried upon dimerization is from the side chains of these residues.

Stability. The leucine zipper coiled coil is stabilized because of three factors: the hydrophobic groups that are buried (leucines at position d and hydrophobic or neutral residues at position a); constancy of size of the internally packing residues at each position; and several distinct ion pairs. Three such ion pairs appear to form, and each is between the e of one heptad and the g of the other. The leucine residues are critical for function in Gcn4. Although each individual leucine can itself be replaced by several different hydrophobic residues, randomized substitution of the leucines with other hydrophobic residues invariably causes the protein to lose function when more than one leucine is substituted; furthermore, isoleucin is by far the most easily tolerated substitution (98).

The binding constant of leucine zipper moieties that interact is estimated to be in the nanomolar rang (163) and has been measured at 5×10^{-8} M for the Jun-Jun dimer at 4°C (196). Even a peptide corresponding to the Fos leucine zipper, which does not dimerize in vitro, has been shown to dimerize in the micromolar range (163).

The leucine zipper moieties that naturally interact do not necessarily have the maximal stability. For exampl , th Gcn4 dimer has a buried asparagine residue which is present within the hydrophobic core (59, 161). This Asn residue packs loosely in the crystal structure, and this position is particularly tolerant of other amino acids (98). Moreover, the asparagine residue (and resultant internal hydrogen bond) drastically destabilizes the Gcn4 zipper; its replacement with valine stabilizes the coiled coil about 1,000-fold (28). It has been speculated that the internal asparagine of Gcn4 (and, by extension, other buried polar groups in the a position in other leucine zippers) is present, so that the proteins do not bind too tightly and therefore can be subject to regulation, or that it keeps the coiled coils in register (4).

Specificity. The specificity of leucine zippers is the key to their regulatory properties. The oncoproteins Fos and Jun, for example, associate with each other to form a heterodimer in preference to the Jun-Jun homodimer. This preference has important consequences in that Fos-Jun heterodimers and Jun-Jun homodimers bend DNA in opposite orientations (114), which may explain the fact that Jun interaction with the glucocorticoid response element of the prolactin gen results in activation of the gene, whereas Fos-Jun interaction results in repression (51).

Specificity of Fos-Jun and Jun-Jun dimerization is achieved primarily by the electrostatic interactions of residues at the e and g positions at the periphery of the hydrophobic core (162). Fos has Glu residues at the g position, and Fos-Fos dimers are much more stable (as measured by T_m) at pH values at which these Glu residues are neutralized. Conversely, Jun is slightly more basic at the e and g positions, and Jun-Jun dimers are more stable at higher pH. Fos-Jun dimers, which are the preferential form, are uniformly stable over a wide range of pH values, because they are more neutral overall. A series of hybrid peptides in an otherwise Gcn4 peptide illustrate the point (162). Specificity (or antispecificity) is achieved by the 8 amino acids at the e and g positions of the peptide and not at other positions.

Regulation. Leucine zipper proteins are likely to be functionally regulated. Thus, the carboxyl-terminal zipper of the human and *Drosophila* heat shock factors may suppress formation of amino-terminal zippers in a way that is sensitive to heat shock (175). Similarly, the calphotin protein binds calcium at one end and has a distinctive leucine zipper at the other end (8). It may therefore be used to transmit signals by altering binding properties.

SH2 Domain

The SH2 domain was first recognized as a noncatalytic domain of Src that was homologous to th Fps protein (189) and is now recognized as a common motif involved in protein-protein interactions (117, 168). More than 20 SH2-containing proteins have been identified. They share a motif of about 100 amino acids that is involved in the recognition of proteins and

BEST AVAILABLE COPY

peptides containing phosphorylated tyrosines. This recognition is implicated in the mechanism of signal transduction, because the phosphorylated tyrosines that are recognized include those of growth factor receptors such as the platelet-derived growth factor receptor, the EGF receptor, and the fibroblast growth factor receptor. On binding their respective growth factors, the growth factor receptors have their tyrosine kinase activity activated, which allows them to autophosphorylate. The autophosphorylated receptor then binds various proteins containing SH2 domains, which are then phosphorylated to modulate their activity. Thus, the binding of growth factor on the outside of the cell results in phosphorylation on the inside of specific substrate proteins. The particular proteins that are phosphorylated depend on the binding specificity of the SH2 domains for the phosphorylated receptor. Binding of different peptides to different SH2 domains has yielded the following results.

Binding of SH2 proteins requires a large domain of the SH2 protein. The conserved domain of SH2 domains, which is common to more than two dozen proteins, has been crystallized for Src (224, 225) and solved by nuclear magnetic resonance spectroscopy techniques for c-Abl (165) and p85a of PI3K (20). In each case, this domain folds into a structure in which a set of internal antiparallel sheets is surrounded by two more or less symmetrical α-helices. The conserved amino acids tend to be part of the recognition for phosphotyrosine (e.g., Arg-175 of Src) or part of the hydrophobic pocket. Variable regions are responsible for sequence recognition (205) and may be parts of variable loops of unknown function (188).

Binding of SH2 proteins requires phosphorylated tyrosine in vitro. Thus, the binding constant of a peptide to an SH2 protein of p85 is between 50- and 800-fold weaker without the phosphate than with the phosphate (64). This preference is attributable to specific side chain contacts of the SH2 domain with the phosphoryl group of phosphotyrosine. The phosphoryl oxygens are hydrogen bonded with two guanidinium hydrogens, one from one arginine and one from another arginine, the hydroxyl hydrogen from threonine and one from serine, and a backbone amide hydrogen. One of the arginines appears to be acting both as a hydrogen bond donor and as an ion pair with the phosphate group. Thus, it cannot be substituted with lysine without loss of binding (140). These contacts are the same whether a weak-affinity (224) or a strong-affinity (225) phosphotyrosine-containing peptide is used.

SH2 domains make contacts with only a small region surrounding the phosphorylated tyrosine. Small peptides faithfully reproduce binding to SH2 domains and display binding constants of the order of nanomolar (64, 218). This is consistent with the crystallographic data of the SH2 domain of v-Src bound to a high-affinity 11-amino-acid peptide; the data clearly show significant peptide-protein interactions at 6 of the 11 positions of the phosphopeptide, from −2 to +3, relative to the tyrosine residue (225). These are the residues that have associated high electron density, indicating a fixed position in the crystal (except for the side chain portion of Gln-1). In addition to the phosphotyrosine-binding interactions described above, there are several ring interactions that define the rest of the phosphotyrosine pocket. There is also a very well-defined interaction of isoleucine at +3 with a deep pocket in the SH2 domain that results in protection of 95% of the surface of the amino acid side chain. The two glutamate residues at +1 and +2 are on the surface of the protein and largely exposed to solvent. Glu+1 appears to interact through its carboxyl group with a lysine amino group, and Glu+2 appears to be stabilized by a nearby arginine guanidinium and its associated H_2O molecules. The amino acids at positions −1 and −2 appear to cap the phosphotyrosine binding through the polypeptide backbone at position −1 and the proline ring at −2.

Other SH2 domain proteins bind different peptides through interactions at the same +1 to +3 positions relative to the phosphotyrosine. This has been elegantly investigated by Songyang et al. (205) through a study of selectivity of binding of random peptides to different SH2 domains. Although the results obtained in this experiment represent bulk selectivity for certain amino acids at certain positions relative to phosphotyrosine, rather than selectivity of individual peptides of known sequence, the results are clear. Each of the three positions following the phosphotyrosine plays an important role in determining the selectivity of binding in certain SH2 proteins, but the amino acids that are crucial and the extent to which they are crucial differs markedly. Thus, most of the discrimination of the C-terminal SH2 domain of p85 is due to its preference for methionine at +3, whereas most of the discrimination of Nck is at positions 1 and 2, where it prefers glutamate and aspartate, respectively (205).

SH3 Domain

The SH3 domain is a second noncatalytic domain of Src which is involved in protein-protein interactions and which is part of a motif shared by other proteins, including tyrosine kinases, phospholipase C-γ (PLC-γ) PI3K, GTPase-activating protein, the cell proliferation proteins Crk and Grb2/Sem5, and the cytoskeletal proteins spectrin, myosin 1, and an actin-binding protein (see references 117, 120, 154, and 168 for a recent list). More than 27 proteins have been shown to have an SH3 domain, which varies between about 55 and 75 amino acids, and its structure has been determined from four different specific domains: spectrin (154), Src (245), PI3K (120), and PLC (118). Each such structure is composed of antiparallel sheets oriented more or less at right angles to one another (or, for PLC, two partial greek key motifs of a barrel oriented such that the strands on opposite sides cross almost perpendicularly), and the amino acids in the conserved strands and a conserved C-terminal 3_{10} helix correspond to many of those that are conserved among SH3 proteins. In each case, a hydrophobic pocket is formed on the surface of the molecule; those of PI3K and Src are remarkably similar (120), and the location of the pocket is conserved between PLC and spectrin (118). This hydrophobic pocket has been implicated in peptide binding for Src (245), since binding of such a peptide perturbs the signal from these amino acids. There are notable differences among the protein structures; PLC, for example, is very similar in secondary structure to spectrin but not to Src, leading to different architectures (118). This property presumably leads to different binding specificities.

The substrates to which SH3-containing proteins bind include an uncharacterized protein similar to GTPase-activating protein-rho, detected with Abl (36); mSos1 and hSos1 (proteins similar to *Drosophila* Sos, which is required for Ras signaling), detected with Grb2 (187); formin and the rat m4 muscarinic receptor, detected with Abl (181); PI3K, detected with v-Src (130); and $p56^{lck}$ and $p59^{fyn}$ (172, 173).

Like the SH2 domain, the SH3 domain binds simple peptides with a high degree of sequence specificity and a high affinity. As judged on a qualitative basis, a 10-amino-acid proline-rich sequence is responsible for strong binding of the Abl SH3 domain to two proteins, called 3BP-1 and 3BP-2 (36, 181). This binding is specific in two ways. First, some but not all single-amino-acid alterations destroy detectable binding. Thus, prolines at positions 2, 7, and 10 are important but those at 5 and to some extent 9 are not. Nonproline residues do not

BEST AVAILABLE COPY

appear to be as important, except perhaps at position 1 (181). Second, peptide binding is SH3 domain specific. Thus, 3BP1 binds the SH3 domains of Abl and Src but not those of Neural Src or Crk (36), and 3BP2 binds most strongly to Abl SH3, less so to Src SH3 and Grb2, and poorly to Nck (181).

Similarly, binding of mSos1 to Grb2 appears to be through a proline-rich motif at the C terminus of the prot in (187); any of several proline-rich 11-amino-acid peptides corresponding to sequences in this region all compete, and competition appears to require a C-terminal arginine. This arginine may add selectivity to the binding of mSos1 to Grb2. A peptide containing the relevant arginine-containing motif binds to Grb2 through its SH3 domain with a K_d of 25 nM (128).

CONCLUDING REMARKS

Alberts and Miake-Lye (5), summarizing a meeting entitled Proteins as Machines, described Tom Pollard's flow diagram for the detailed analysis of a cell biology process. First, a complete inventory of all the molecules making up the machine must be made. Second, a determination must be made of how and in what order the molecules interact with each other. Third, both detailed rate constants for each transition and structures of each component at atomic resolution must be obtained. While no process is yet completely understood at the three levels described by Pollard, enormous progress has been made in deciphering protein machines. In this review, we have tried to convey some of the classical and more recent approaches used to develop the inventory of proteins and the nature of their interactions.

Two factors are having a large impact on how cellular processes are viewed. First, the vast amount of DNA sequence information being obtained means that the identity of almost all proteins, at the level of primary sequence, may soon be known. Complete sequences for organisms such as *E. coli*, yeast cells, and the nematodes and nearly complete compilations of the cDNA sequences for human tissues should be available in the next few years. Second, the range of new procedures now available means that hundreds to thousands of new protein-protein interactions may be identified in the same period. Ten to twenty years ago, only a few complexes of proteins were well characterized as to their subunit composition and specific interactions; currently, a large number of such complexes are known. Relatively soon, there may be an enormous number. The continuing challenge will be for biochemists and cell, molecular, and structural biologists to use this information to understand how the cell works.

ACKNOWLEDGMENTS

We thank R. Angerer, M. Dumont, E. Grayhack, T. Miller, F. Sherman, and L. Tabak for many helpful comments on the manuscript and J. Nimmo for superb secretarial assistance.

Research from our laboratories has been supported by grants MCB-9118571 from the National Science Foundation and VM-99 from the American Cancer Society (to E.M.P.) and grants GM 49065 and CA 28146 from the National Institutes of Health (to S.F.).

REFERENCES

1. Ackers, G. K. 1970. Analytical gel chromatography of proteins. Adv. Protein Chem. 24:343–441.
2. Ackers, G. K. 1975. Molecular sieve methods of analysis, p. 1–94. *In* H. Neurath and R. L. Hill (ed.), The proteins. Academic Press, Inc., New York.
3. Adams, A. E. M., and D. Botstein. 1989. Dominant suppressors of yeast actin mutations that are reciprocally suppressed. Genetics 121:675–683.
4. Alber, T. 1992. Structure of the leucine zipper. Curr. Opin. Genet. Dev. 2:205–210.
5. Alberts, B., and R. Miake-Lye. 1992. Unscrambling the puzzle of biological machines: the importance of the details. Cell 68:415–420.
6. Ayer, D. E., L. Kretzner, and R. N. Eisenman. 1993. Mad: a heterodimeric partner for Max that antagonizes Myc transcriptional activity. Cell 72:211–222.
7. Baird, B. A., and G. G. Hammes. 1976. Chemical cross-linking studies of chloroplast coupling factor 1. J. Biol. Chem. 251:6953–6962.
8. Ballinger, D. G., N. Xue, and K. D. Harshman. 1993. A *Drosophila* photoreceptor cell-specific protein, calphotin, binds calcium and contains a leucine zipper. Proc. Natl. Acad. Sci. USA 90:1536–1540.
9. Banik, U., N. C. Mandal, B. Bhattacharyya, and S. Roy. 1993. A fluorescence anisotropy study of tetramer-dimer equilibrium of λ repressor and its implication for function. J. Biol. Chem. 268:3938–3943.
10. Barbas, C. F., A. S. Kang, R. A. Lerner, and S. J. Benkovic. 1991. Assembly of combinatorial antibody libraries on phage surfaces: the gene III site. Proc. Natl. Acad. Sci. USA 88:7978–7982.
11. Bass, S., R. Greene, and J. A. Wells. 1990. Hormone phage: an enrichment method for variant proteins with altered binding properties. Proteins 8:309–314.
12. Basson, M. E., R. L. Moore, J. O'Rear, and J. Rine. 1987. Identifying mutations in duplicated functions in *Saccharomyces cerevisiae* by use of the *LYS2* gene. Genetics 117:645–655.
13. Beckett, D., K. S. Koblan, and G. K. Ackers. 1991. Quantitative study of protein association at picomolar concentrations: the λ phage cI repressor. Anal. Biochem. 196:69–75.
14. Beeckmans, S., and L. Kanarek. 1981. Demonstration of physical interactions between consecutive enzymes of the citric acid cycle and of the aspartate-malate shuttle. Eur. J. Biochem. 117:527–535.
15. Bender, T., and J. R. Pringle. 1991. Use of a screen for synthetic lethal and multicopy suppressee mutants to identify two new genes involved in morphogenesis in *Saccharomyces cerevisiae*. Mol. Cell. Biol. 11:1295–1305.
16. Birge, R. B., J. E. Fajardo, C. Reichman, S. E. Shoelson, Z. Songyang, L. C. Cantley, and H. Hanafusa. 1993. Identification and characterization of high-affinity interaction between v-Crk and tyrosine-phosphorylated paxillin in CT10-transformed fibroblasts. Mol. Cell. Biol. 13:4648–4656.
17. Blackwood, E. M., and R. N. Eisenman. 1991. Max: a helix-loop-helix zipper protein that forms a sequence-specific DNA-binding complex with myc. Science 251:1211–1217.
18. Blanar, M. A., and W. J. Rutter. 1992. Interaction cloning: Identification of a helix-loop-helix zipper protein that interacts with c-Fos. Science 256: 1014–1018.
19. Bleil, J. D., and P. M. Wassarman. 1990. Identification of a ZP3-binding protein on acrosome-intact mouse sperm by photoaffinity crosslinking. Proc. Natl. Acad. Sci. USA 87:5563–5567.
20. Booker, G. W., A. L. Breeze, A. K. Downing, G. Panayotou, I. Gout, M. D. Waterfield, and I. D. Campbell. 1992. Structure of an SH2 domain of the p85α subunit of phosphatidylinositol-3-OH kinase. Nature (London) 358: 684–687.
21. Bragg, P. D., and C. Hou. 1980. A cross-linking study of the Ca^{2+}, Mg^{2+}-activated adenosine triphosphatase of *Escherichia coli*. Eur. J. Biochem. 106:495–503.
22. Bragg, P. D., and C. Hou. 1986. Chemical crosslinking of α subunits in the F_1 adenosine triphosphatase of *Escherichia coli*. Arch. Biochem. Biophys. 244:361–372.
23. Brent, R., and M. Ptashne. 1985. A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor. Cell 43:729–736.
24. Bretscher, A., and K. Weber. 1980. Villin is a major protein of the microvillus cytoskeleton which binds both G and F actin in a calcium-dependent manner. Cell 20:839–847.
25. Brown, R. L. 1992. Functional regions of the inhibitory subunit of retinal rod cGMP phosphodiesterase identified by site-specific mutagenesis and fluorescence spectroscopy. Biochemistry 31:5918–5925.
26. Burton, D. R., C. F. Barbas, M. A. A. Persson, S. Koenig, R. M. Chanock, and R. A. Lerner. 1991. A large array of human monoclonal antibodies to type 1 human immunodeficiency virus from combinatorial libraries of asymptomatic seropositive individuals. Proc. Natl. Acad. Sci. USA 88:10134–10137.
27. Calakos, N., M. K. Bennett, K. E. Peterson, and R. H. Scheller. 1994. Protein-protein interactions contributing to the specificity of intracellular vesicular trafficking. Science 263:1146–1149.
28. Carr, C., J. Petersen, P. S. Kim, and T. Alber. 1992. Unpublished data.
29. Carr, D. W., and J. D. Scott. 1992. Blotting and band-shifting: techniques for studying protein-protein interactions. Trends Biochem. Sci. 17:246–249.
30. Chardin, P., J. H. Camonis, N. W. Gale, L. V. Aelst, J. Schlessinger, M. H. Wigler, and D. Bar-Sagi. 1993. Human Sos1: a guanine nucleotide exchange factor for Ras that binds to GRB2. Science 260:1338–1343.
31. Chellappan, S., V. B. Kraus, B. Kroger, K. Munger, P. M. Howley, W. C. Phelps, and J. R. Nevins. 1992. Adenovirus E1A, simian virus 40 tumor antigen, and human papillomavirus E7 protein share the capacity to disrupt the interaction between transcription factor E2F and the retinoblastoma gene product. Proc. Natl. Acad. Sci. USA 89:4549–4553.
32. Chellappan, S. P., S. Hiebert, M. Mudryj, J. M. Horowitz, and J. R. Nevins.

BEST AVAILABLE COPY

1991. The E2F transcription factor is a cellular target for the RB protein. Cell 65:1053–1061.
Chen, Y., Y. W. Ebright, and R. H. Ebright. 1994. Identification of the target of a transcription activator protein by protein-protein photocrosslinking. Science 265:90–92.
34. Chevray, P. M., and D. Nathans. 1992. Protein interaction cloning in yeast: identification of mammalian proteins that react with the leucine zipper of Jun. Proc. Natl. Acad. Sci. USA 89:5789–5793.
Chien, C.-T., P. L. Bartel, R. Sternglanz, and S. Fields. 1991. The two-hybrid system: a method to identify and clone genes for proteins that interact with a protein of interest. Proc. Natl. Acad. Sci. USA 88:9578–9582.
Cicchetti, P., B. J. Mayer, G. Thiel, and D. Baltimore. 1992. Identification of a protein that binds to the SH3 region of Abl and is similar to Bcr and GAP-rho. Science 257:803–806.
37. Clackson, T., H. R. Hoogenboom, A. D. Griffiths, and G. Winter. 1991. Making antibody fragments using phage display libraries. Nature (London) 352:624–628.
Cohen, B. D., D. R. Lowy, and J. T. Schiller. 1993. The conserved C-terminal domain of the bovine papillomavirus E5 oncoprotein can associate with an α-adaptin-like molecule: a possible link between growth factor receptors and viral transformation. Mol. Cell. Biol. 13:6462–6468.
Cornish, V. W., D. R. Benson, C. A. Altenbach, K. Hideg, W. L. Hubbell, and P. G. Schultz. 1994. Site-specific incorporation of biophysical probes into proteins. Proc. Natl. Acad. Sci. USA 91:2910–2914.
40. Costigan, C., S. Gehrung, and M. Snyder. 1992. A synthetic lethal screen identifies SLK1, a novel protein kinase homolog implicated in yeast cell morphogenesis and cell growth. Mol. Cell. Biol. 12:1162–1178.
Cover, J. A., J. M. Lambert, C. M. Norman, and R. R. Traut. 1981. Identification of proteins at the subunit interface of the *Escherichia coli* ribosome by cross-linking with dimethyl 3,3'-dithiobis(propionimidate). Biochemistry 20:2843–2852.
Crick, F. H. C. 1953. The packing of α-helices: simple coiled-coils. Acta Crystallogr. 6:689–697.
43. Cull, M. G., J. F. Miller, and P. J. Schatz. 1992. Screening for receptor ligands using large libraries of peptides linked to the C terminus of the *lac* repressor. Proc. Natl. Acad. Sci. USA 89:1865–1869.
Cwirla, S. E., E. A. Peters, R. W. Barrett, and W. J. Dower. 1990. Peptides on phage: a vast library of peptides for identifying ligands. Proc. Natl. Acad. Sci. USA 87:6378–6382.
45. Dalton, S., and R. Treisman. 1992. Characterization of SAP-1, a protein recruited by serum response factor to the c-fos serum response element. Cell 68:597–612.
Dang, C. V., J. Barrett, M. Villa-Garcia, L. M. S. Resar, G. J. Kato, and E. R. Fearon. 1991. Intracellular leucine zipper interactions suggest c-Myc hetero-oligomerization. Mol. Cell. Biol. 11:954–962.
47. Darawshe, S., G. Rivas, and A. P. Minton. 1993. Sedimentation equilibrium-quantitative polyacrylamide gel electrophoresis (SE-QPAGE): a new technique for the detection of associations in multicomponent solutions. Anal. Biochem. 215:236–242.
48. de Gunzberg, J., R. Riehl, and R. A. Weinberg. 1989. Identification of a protein associated with $p21^{ras}$ by chemical crosslinking. Proc. Natl. Acad. Sci. USA 86:4007–4011.
Denny, J. B., and B. Blobel. 1984. ^{125}I-labeled crosslinking reagent that is hydrophilic, photoactivatable, and cleavable through an azo linkage. Proc. Natl. Acad. Sci. USA 81:5286–5290.
50. Devlin, J. J., L. C. Panganiban, and P. E. Devlin. 1990. Random peptide libraries: a source of specific protein binding molecules. Science 249:404–406.
51. Diamond, M. I., J. N. Miner, S. K. Yoshinaga, and K. R. Yamamoto. 1990. Transcription factor interactions: selectors of positive or negative regulation from a single DNA element. Science 249:1266–1272.
Dobzhansky, T. 1946. Genetics of natural populations. XIII. Recombination and variability in populations of *Drosophila pseudoobscura*. Genetics 31:269–290.
53. Drubin, D. G., K. G. Miller, and D. Botstein. 1988. Yeast actin-binding proteins: evidence for a role in morphogenesis. J. Cell Biol. 107:2551–2561.
Dumont, M. E., A. F. Corin, and G. A. Campbell. 1994. Noncovalent binding of heme induces a compact apocytochrome *c* structure. Biochemistry 33:7368–7378.
55. Durfee, T., K. Becherer, R.-L. Chen, S. H. Yeh, Y. Yang, A. E. Kilburn, W. H. Lee, and S. J. Elledge. 1993. The retinoblastoma protein associates with the protein phosphatase type 1 catalytic subunit. Genes Dev. 7:555–569.
56. Durfee, T., B. Gorovits, C. Hensey, and W.-H. Lee. *In vivo* and *in vitro* determined binding affinities of complexes involving the retinoblastoma protein exhibit an exponential relationship. Submitted for publication.
Eisenstein, E., and H. K. Schachman. 1989. Determining the roles of subunits in protein function, p. 135–175. *In* T. E. Creighton (ed.), Protein function: a practical approach. IRL Press, Oxford.
58. El-Deiry, W. S., T. Tokino, V. E. Velculescu, D. B. Levy, R. Parson, J. M. Trent, D. Lin, W. E. Mercer, K. W. Kinzler, and B. Vogelstein. 1993. *WAF1*, a potential mediator of p53 tumor suppression. Cell 75:817–825.
59. Ellenberger, T. E., C. J. Brandl, K. Struhl, and S. C. Harrison. 1992. The GCN4 basic region leucine zipper binds DNA as a dimer of uninterrupted α helices: crystal structure of the protein-DNA complex. Cell 71:1226–1237.
60. Evans, P. R., G. W. Farrants, and P. J. Hudson. 1981. Phosphofructokinase: structure and control. Philos. Trans. R. Soc. London 293:53–62.
61. Ewen, M. E., B. Faha, E. Harlow, and D. M. Livingston. 1992. Interaction of p107 with cyclin A independent of complex formation with viral oncoproteins. Science 255:85–87.
62. Faha, B., M. E. Ewen, L.-H. Tsai, D. M. Livingston, and E. Harlow. 1992. Interaction between human cyclin A and adenovirus E1A-associated p107 protein. Science 255:87–90.
63. Fahien, L. A., and S. E. Smith. 1974. The enzyme-enzyme complex of transaminase and glutamate dehydrogenase. J. Biol. Chem. 249:2696–2703.
64. Felder, S., M. Zhou, P. Hu, J. Urena, A. Ullrich, M. Chaudhuri, S. E. Shoelson, and J. Schlessinger. 1993. SH2 domains exhibit high-affinity binding to tyrosine-phosphorylated peptides yet also exhibit rapid dissociation and exchange. Mol. Cell. Biol. 13:1449–1455.
65. Fields, S., and O.-K. Song. 1989. A novel genetic system to detect protein-protein interactions. Nature (London) 340:245–246.
66. Fields, S., and R. Sternglanz. 1994. The two-hybrid system: An assay for protein-protein interactions. Trends Genet. 10:286–292.
67. Finlay, C. A., P. W. Hinds, and A. J. Levine. 1989. The p53 proto-oncogene can act as a suppressor of transformation. Cell 57:1083–1093.
68. Flynn, D. C., T.-H. Leu, A. B. Reynolds, and J. T. Parson. 1993. Identification and sequence analysis of cDNAs encoding a 110-kilodalton actin filament-associated $pp60^{src}$ substrate. Mol. Cell. Biol. 13:7892–7900.
69. Formosa, T., J. Barry, B. M. Alberts, and J. Greenblatt. 1991. Using protein affinity chromatography to probe structure of protein machines. Methods Enzymol. 208:24–45.
70. Friedrich, P. 1984. Supramolecular enzyme organization: quarternary structure and beyond. Pergamon Press, Oxford.
71. Friguet, B., A. F. Chaffotte, L. Djavadi-Ohaniance, and M. Goldberg. 1985. Measurements of the true affinity constant in solution of antigen-antibody complexes. B1 enzyme-linked immunosorbent assay. J. Immunol. Methods 77:305–319.
72. Gegner, J. A., and F. W. Dahlquist. 1991. Signal transduction in bacteria: CheW forms a reversible complex with the protein kinase CheA. Proc. Natl. Acad. Sci. USA 88:750–754.
73. Geiser, J. R., H. A. Sundberg, B. H. Chang, E. G. D. Muller, and T. N. Davis. 1993. The essential mitotic target of calmodulin is the 110-kilodalton component of the spindle pole body in *Saccharomyces cerevisiae*. Mol. Cell. Biol. 13:7913–7924.
74. Germino, F. J., Z. X. Wang, and S. M. Weissman. 1993. Screening for *in vivo* protein-protein interactions. Proc. Natl. Acad. Sci. USA 90:933–937.
75. Gilbert, G. A., and G. L. Kellett. 1971. Interacting systems of the type A+B=C: ovalbumin and myoclobin. J. Biol. Chem. 246:6079–6086.
76. Gill, S. C., S. E. Weitzel, and P. H. Von Hippel. 1991. *Escherichia coli* σ^{70} and NusA proteins. 1. Binding interactions with core RNA polymerase in solution and within the transcription complex. J. Mol. Biol. 220:307–324.
77. Glenney, J. R., and K. Weber. 1993. Detection of calmodulin-binding polypeptides separated in SDS-polyacrylamide gels by a sensitive [^{125}I]calmodulin gel overlay assay. Methods Enzymol. 102:204–210.
78. Goodsell, D. S., and A. J. Olson. 1993. Soluble proteins: size, shape and function. Trends Biochem. Sci. 18:65–68.
79. Gorlich, D., E. Hartmann, S. Prehn, and T. A. Rapoport. 1992. A protein of the endoplasmic reticulum involved early in polypeptide translocation. Nature (London) 357:47–52.
80. Greenblatt, J., and J. Li. 1981. The *nusA* gene protein of *Escherichia coli*. J. Mol. Biol. 147:11–23.
81. Gross, M., and T. Wallimann. 1993. Kinetics of assembly and dissociation of the mitochondrial creatine kinase octamer: a fluorescence study. Biochemistry 32:13933–13940.
82. Hadwiger, J. A., C. Wittenberg, H. E. Richardson, M. D. Lopes, and S. I. Reed. 1989. A family of cyclin homologs that control the G1 phase in yeast. Proc. Natl. Acad. Sci. USA 86:6255–6259.
83. Hannon, G. J., D. Demetrick, and D. Beach. 1993. Isolation of the Rb-related p130 through its interaction with CDK2 and cyclins. Genes Dev. 7:2378–2391.
84. Harlow, E., P. Whyte, B. R. Franza, and C. Schley. 1986. Association of adenovirus early-region 1A proteins with cellular polypeptides. Mol. Cell. Biol. 6:1579–1589.
85. Harper, J. W., G. R. Adami, N. Wei, K. Keyomarsi, and S. J. Elledge. 1993. The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases. Cell 75:805–816.
86. Hartman, P. E., and J. R. Roth. 1973. Mechanisms of suppression. Adv. Genet. 17:1–105.
87. Hartmann, E., M. Wiedmann, and T. A. Rapoport. 1989. A membrane component of the endoplasmic reticulum that may be essential for protein translocation. EMBO J. 8:2225–2229.
88. Herberg, F. W., W. R. G. Dostmann, M. Zorn, S. J. Davis, and S. S. Taylor. 1994. Crosstalk between domains in the regulatory subunit of cAMP-dependent protein kinase: Influence of amino terminus on cAMP binding and

BEST AVAILABLE COPY

holoenzyme formation. Biochemistry 33:7485–7494.
89. Herberg, F. W., and S. S. Taylor. 1993. Physiological inhibitors of the catalytic subunit of cAMP-dependent protein kinase: effect of MgATP on protein-protein interactions. Biochemistry 32:14015–14022.
90. Herskowitz, I. 1987. Functional inactivation of genes by dominant negative mutations. Nature (London) 329:219–222.
91. Heyduk, T., J. C. Lee, Y. W. Ebright, E. E. Blatter, Y. Zhou, and R. H. Ebright. 1993. CAP interacts with RNA polymerase in solution in the absence of promoter DNA. Nature (London) 364:548–549.
92. Hiebert, S. W., S. P. Chellappan, J. M. Horowitz, and J. R. Nevins. 1992. The interaction of RB with E2F coincides with an inhibition of the transcriptional activity of E2F. Genes Dev. 6:177–185.
93. Hieter, P., C. Mann, M. Snyder, and R. W. Davis. 1985. Mitotic stability of yeast chromosomes: a colony color assay that measures nondisjunction and chromosome loss. Cell 40:381–392.
94. High, S., B. Bartoglio, D. Gorlich, S. S. L. Andersen, A. J. Ashford, A. Giner, E. Hartmann, S. Prehn, T. A. Rapoport, B. Dobberstein, and J. Brunner. 1993. Site-specific photocross-linking reveals that Sec61p and TRAM contact different regions of a membrane-inserted signal sequence. J. Biol. Chem. 368:28745–28751.
95. Hill, R. L., and K. Brew. 1975. Lactose synthetase. Adv. Enzymol. 43:411–490.
96. Hodgkin, J., K. Kondo, and R. H. Waterston. 1987. Suppression in the nematode *Caenorhabditis elegans*. Trends Genet. 3:325–329.
97. Hope, I. A., and K. Struhl. 1986. Functional dissection of a eukaryotic transcriptional activator protein, GCN4 of yeast. Cell 46:885–894.
98. Hu, J. C., E. K. O'Shea, P. S. Kim, and R. T. Sauer. 1990. Sequence requirements for coiled-coils: analysis with λ repressor-GCN4 leucine zipper fusions. Science 250:1400–1403.
99. Hu, P., B. Margolis, E. Y. Skolnik, R. Lammers, A. Ullrich, and J. Schlessinger. 1992. Interaction of phosphatidylinositol 3-kinase-associated p85 with epidermal growth factor and platelet-derived growth factor receptors. Mol. Cell. Biol. 12:981–990.
100. Huang, S., W.-H. Lee, and E. Y.-H. P. Lee. 1991. A cellular protein that competes with SV40 T antigen for binding to the retinoblastoma gene product. Nature (London) 350:160–162.
101. Huffaker, T. C., M. A. Hoyt, and D. Botstein. 1987. Genetic analysis of the yeast cytoskeleton. Annu. Rev. Genet. 21:259–284.
102. Hummel, J. P., and W. J. Dreyer. 1962. Measurement of protein-binding phenomena by gel filtration. Biochim. Biophys. Acta 63:530–532.
103. Hurley, J. B., and L. Stryer. 1992. Purification and characterization of the γ regulatory subunit of the cyclic GMP phosphodiesterase from retinal rod outer segments. J. Biol. Chem. 257:11094–11099.
104. Il'ichev, A. A., O. O. Minenkova, S. I. Tatkov, N. N. Karpyshev, A. M. Eroshkin, V. A. Petrenko, and L. S. Sandakhchiev. 1989. M13 filamentous bacteriophage in protein engineering. Dokl. Akad. Nauk SSSR 307:431–433.
105. Iwabuchi, K., B. Li, P. L. Bartel, and S. Fields. 1993. Use of the two-hybrid system to identify the domain of p53 involved in oligomerization. Oncogene 8:1693–1696.
106. James, P., M. Inui, M. Tada, M. Chiesi, and E. Carafoli. 1989. Nature and site of phospholamban regulation of the Ca^{2+} pump of sarcoplasmic reticulum. Nature (London) 342:90–92.
107. Jarvik, J., and D. Botstein. 1975. Conditional-lethal mutations that suppress genetic defects in morphogenesis by altering structural proteins. Proc. Natl. Acad. Sci. USA 72:2738–2742.
108. Jarvis, T. C., D. M. Ring, S. S. Daube, and P. H. Von Hippel. 1990. "Macromolecular crowding": thermodynamic consequences for protein-protein interactions within the T4 DNA replication complex. J. Biol. Chem. 265:15160–15167.
109. Jonsson, U., L. Fagerstam, B. Ivarsson, B. Johnsson, R. Karlsson, K. Lundh, S. Lofas, B. Persson, H. Roos, I. Ronnberg, S. Sjolander, E. Stenberg, R. Stahlberg, C. Urbaniczky, H. Ostlin, and M. Malmqvist. 1991. Real-time biospecific interaction analysis using surface plasmon resonance and a sensor chip technology. BioTechniques. 11:620–627.
110. Kang, A. S., C. F. Barbas, K. D. Janda, S. J. Benkovic, and R. A. Lerner. 1991. Linkage of recognition and replication functions by assembling combinatorial antibody Fab libraries along phage surfaces. Proc. Natl. Acad. Sci. USA 88:4363–4366.
111. Karlsson, R., A. Michaelsson, and L. Mattsson. 1991. Kinetic analysis of monoclonal antibody-antigen interactions with a new biosensor based analytical system. J. Immunol. Methods 145:229–240.
112. Keegan, L., G. Gill, and M. Ptashne. 1986. Separation of DNA binding from the transcription-activating function of a eukaryotic regulatory protein. Science 231:699–704.
113. Kellogg, D. R., C. M. Field, and B. M. Alberts. 1989. Identification of microtubule-associated proteins in the centrosome, spindle, and kinetochore of the early *Drosophila* embryo. J. Cell Biol. 109:2977–2991.
114. Kerppola, T. K., and T. Curran. 1991. Fos-Jun heterodimers and Jun homodimers bend DNA in opposite orientations: implications for transcription factor cooperativity. Cell 66:317–326.
115. Kim, Y. T., S. Tabor, J. E. Churchich, and C. C. Richardson. 1992. Interactions of gene 2.5 protein and DNA polymerase of bacteriophage T7. J. Biol. Chem. 267:15032–15040.
116. Klotz, I. M., D. W. Darnall, and N. R. Langerman. 1975. Quarternary structure of proteins, p. 293–411. *In* H. Neurath and R. L. Hill (ed.), The proteins. Academic Press, Inc., New York.
117. Koch, C. A., D. Anderson, M. F. Moran, C. Willis, and T. Pawson. 1991. SH2 and SH3 domains: elements that control interactions of cytoplasmic signaling proteins. Science 252:668–674.
118. Kohda, D., H. Hatanaka, M. Odaka, V. Mandiyan, A. Ullrich, J. Schlessinger, and F. I. Nagaki. 1993. Solution structure of the SH3 domain of phospholipase C-g. Cell 72:953–960.
119. Kosbland, D., J. C. Kent, and L. H. Hartwell. 1985. Genetic analysis of the mitotic transmission of minochromosomes. Cell 40:393–403.
120. Koyama, S., H. Yu, D. C. Dalgarno, T. B. Shin, L. D. Zydowsky, and S. L. Schreiber. 1993. Structure of the PI3K SH3 domain and analysis of the SH3 family. Cell 72:945–952.
121. Krieg, U. C., A. E. Johnson, and P. Walter. 1989. Protein translocation across the endoplasmic reticulum membrane: identification by photocross-linking of a 39-kD integral membrane glycoprotein as part of a putative translocation tunnel. J. Cell Biol. 109:2033–2043.
122. Kukuruzinska, M. A., B. W. Turner, G. K. Ackers, and S. Roseman. 1984. Subunit association of enzyme I of the *Salmonella typhimurium* phosphoenolpyruvate:glycose phosphotransferase system. J. Biol. Chem. 259: 11679–11681.
123. Kwon, O.-S., and J. E. Churchich. 1994. Interaction of 70-kDa heat shock cognate protein with peptides and *myo*-inositol monophosphatase. J. Biol. Chem. 269:266–271.
124. Landschulz, W. H., P. F. Johnson, and S. L. McKnight. 1988. The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. Science 240:1759–1764.
125. Lazarides, E., and U. Lindberg. 1974. Actin is the naturally occurring inhibitor of deoxyribonuclease I. Proc. Natl. Acad. Sci. USA 71:4742–4746.
126. Lee, F. S., D. S. Auld, and B. L. Vallee. 1989. Tryptophan fluorescence as a probe of placental ribonuclease inhibitor binding to angiogenin. Biochemistry 28:219–224.
126a. Lee, F. S., R. Shapiro, and B. L. Vallee. 1989. Tight-binding inhibition of angiogenin and ribonuclease A by placental ribonuclease inhibitor. Biochemistry 28:225–230.
127. Li, B., and S. Fields. 1993. Identification of mutations in p53 that affect its binding to SV40 large T antigen by using the yeast two-hybrid system. FASEB J. 7:957–963.
128. Li, N., A. Batzer, R. Daly, V. Yajnik, E. Skolnik, P. Chardin, D. Bar-Sagi, B. Margolis, and J. Schlessinger. 1993. Guanine-nucleotide-releasing factor hSos1 binds to Grb2 and links receptor tyrosine kinases to Ras signalling. Nature (London) 363:85–88.
129. Li, S., and J. M. Sedivy. 1993. Raf-1 protein kinase activates the NF-κB transcription factor by dissociating the cytoplasmic NF-kB-IkB complex. Proc. Natl. Acad. Sci. USA 90:9247–9251.
130. Liu, X., L. E. M. Marengere, C. A. Koch, and T. Pawson. 1993. The v-Src SH3 domain binds phosphatidylinositol 3'-kinase. Mol. Cell. Biol. 13:5225–5232.
131. Lowman, H. B., S. H. Bass, N. Simpson, and J. A. Wells. 1991. Selecting high-affinity binding proteins by monovalent phage display. Biochemistry 30:10832–10838.
132. Ludlow, J. W., J. A. DeCaprio, C.-M. Huang, W.-H. Lee, E. Paucha, and D. M. Livingston. 1989. SV40 large T antigen binds preferentially to an underphosphorylated member of the retinoblastoma susceptibility gene product family. Cell 56:57–65.
133. Ludlow, J. W., C. L. Glendening, D. M. Livingston, and J. A. DeCaprio. 1993. Specific enzymatic dephosphorylation of the retinoblastoma protein. Mol. Cell. Biol. 13:367–372.
134. Ma, J., and M. Ptashne. 1987. A new class of yeast transcriptional activators. Cell 51:113–119.
135. MacGregor, P. F., C. Abate, and T. Curran. 1990. Direct cloning of leucine zipper proteins: jun binds cooperatively to the CRE with CRE-BP1. Oncogene 5:451–458.
136. Malmqvist, M. 1993. Biospecific interaction analysis using biosensor technology. Nature (London) 361:186–187.
137. Marengere, L. E. M., Z. Songyang, G. D. Gish, M. D. Schaller, J. T. Parsons, M. J. Stern, L. C. Cantley, and T. Pawson. 1994. SH2 domain specificity and activity modified by a single residue. Nature (London) 369: 502–505.
138. Mason, S. W., J. Li, and J. Greenblatt. 1992. Direct interaction between two *Escherichia coli* transcription antitermination factors, NusB and ribosomal protein S10. J. Mol. Biol. 223:55–66.
139. Mayer, B. J., P. K. Jackson, and D. Baltimore. 1991. The noncatalytic src homology region 2 segment of *abl* tyrosine kinase binds to tyrosine-phosphorylated cellular proteins with high affinity. Proc. Natl. Acad. Sci. USA 88:627–631.
140. Mayer, B. J., P. K. Jackson, R. A. Van Etten, and D. Baltimore. 1992. Point mutations in the *abl* SH2 domain coordinately impair phosphotyrosine binding in vitro and transforming activity in vivo. Mol. Cell. Biol. 12:609–618.

BEST AVAILABLE COPY

McCafferty, J., A. D. Griffiths, G. Winter, and D. J. Chiswell. 1990. Phage antibodies: filamentous phage displaying antibody variable domains. Nature (London) 348:552–554.

Meeks-Wagner, D., and L. H. Hartwell. 1986. Normal stoichiometry of histone dimer sets is necessary for high fidelity of mitotic chromosome transmission. Cell 44:43–52.

Miller, J. H. 1978. The *lacI* gene: its role in *lac* operon control and its use as a genetic system, p. 31–88. *In* J. H. Miller and W. S. Reznikoff (ed.), The operon. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Miller, K. G., and B. M. Alberts. 1989. F-actin affinity chromatography: technique for isolating previously unidentified actin-binding proteins. Proc. Natl. Acad. Sci. USA 86:4808–4812.

Miller, K. G., C. M. Field, B. M. Alberts, and D. R. Kellogg. 1991. Use of actin filament and microtubule affinity chromatography to identify proteins that bind to the cytoskeleton. Methods Enzymol. 196:303–319.

Mills, J. S., M. P. Walsh, K. Nemcek, and J. D. Johnson. 1988. Biologically active fluorescent derivatives of spinach calmodulin that report calmodulin target protein binding. Biochemistry 27:991–996.

Mita, S., A. Tominaga, Y. Hitoshi, K. Sakamoto, T. Honjo, M. Akagi, Y. Kikuchi, N. Yamaguchi, and K. Takatsu. 1989. Characterization of high-affinity receptors for interleukin 5 on interleukin 5-dependent cell lines. Proc. Natl. Acad. Sci. USA 86:2311–2315.

Mitchell, D. A., T. K. Marshall, and R. J. Deschenes. 1993. Vectors for the inducible overexpression of glutathione S-transferase fusion proteins in yeast. Yeast 9:715–723.

Mitsuzawa, H., I. Uno, T. Oshima, and T. Ishikawa. 1989. Isolation and characterization of temperature-sensitive mutations in the *RAS2* and *CYR1* genes of *Saccharomyces cerevisiae*. Genetics 123:739–748.

Moerman, D. G., G. M. Benian, R. J. Barstead, L. A. Schriefer, and R. H. Waterston. 1988. Identification and intracellular localization of the *unc-22* gene product of *Caenorhabditis elegans*. Genes Dev. 2:93–105.

Moerman, D. G., S. Plurad, and R. H. Waterston. 1982. Mutations in the *unc-54* myosin heavy chain gene of *Caenorhabditis elegans* that alter contractility but not muscle structure. Cell 29:773–781.

Moir, D., S. E. Stewart, B. C. Osmond, and D. Botstein. 1982. Cold-sensitive cell-division-cycle mutants of yeast: isolation, properties, and pseudoreversion studies. Genetics 100:547–563.

Monod, J., J. Wyman, and J.-P. Changeux. 1965. On the nature of allosteric transitions: a plausible model. J. Mol. Biol. 12:88–118.

Musacchio, A., M. Noble, R. Pauptit, R. Wierenga, and M. Saraste. 1992. Crystal structure of a Src-homology 3 (SH3) domain. Nature (London) 359:851–855.

Musch, A., M. Wiedmann, and T. A. Rapoport. 1992. Yeast Sec proteins interact with polypeptides traversing the endoplasmic reticulum membrane. Cell 69:343–352.

Nakai, H., and C. C. Richardson. 1990. The gene 1.2 protein of bacteriophage T7 interacts with the *Escherichia coli* dGTP triphosphohydrolase to form a GTP-binding protein. J. Biol. Chem. 265:4411–4419.

Nefsky, B., and A. Bretscher. 1992. Yeast actin is relatively well behaved. Eur. J. Biochem. 206:949–955.

Nelson, R. M., and G. L. Long. 1991. Solution-phase equilibrium binding interaction of human protein S with C4b-binding protein. Biochemistry 30:2384–2390.

Nogi, Y., K. Matsumoto, A. Toh-e, and Y. Oshima. 1977. Regulation of expression of the galactose gene cluster in *Saccharomyces cerevisiae*. II. The isolation and dosage effect of the regulatory gene Gal80. Mol. Gen. Genet. 195:29–34.

Novick, P., B. C. Osmond, and D. Botstein. 1989. Suppressors of yeast actin mutations. Genetics 121:659–674.

O'Shea, E. K., J. D. Klemm, P. S. Kim, and T. Alber. 1991. X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. Science 254:539–544.

O'Shea, E. K., R. Rutkowski, and P. S. Kim. 1992. Mechanism of specificity in the Fos-Jun oncoprotein heterodimer. Cell 68:699–708.

O'Shea, E. K., R. Rutkowski, W. F. Stafford III, and P. S. Kim. 1989. Preferential heterodimer formation by isolated leucine zippers from fos and jun. Science 245:646–648.

Otto-Bruc, A., B. Antonny, T. M. Vuong, P. Chardin, and M. Chabre. 1993. Interaction between retinal cyclic GMP phosphodiesterase inhibitor and transducin: kinetics and affinity studies. Biochemistry 32:8636–8645.

Overduin, M., C. B. Rios, B. J. Mayer, D. Baltimore, and D. Cowburn. 1992. Three dimensional solution structure of the scr homology 2 domain of c-abl. Cell 70:697–704.

Panayotou, G., G. Gish, P. End, O. Truong, I. Gout, R. Dhand, M. J. Fry, I. Hiles, T. Pawson, and M. D. Waterfield. 1993. Interactions between SH2 domains and tyrosine-phosphorylated platelet-derived growth factor β-receptor sequences: analysis of kinetic parameters by a novel biosensor-based approach. Mol. Cell. Biol. 13:3567–3576.

Pannekoek, H., M. VanMeijer, R. R. Schleef, D. J. Loskutoff, and C. F. Barbas. 1993. Functional display of human plasminogen-activator inhibitor 1 (PAI-1) on phages: novel perspectives for structure-function analysis by error-prone DNA synthesis. Gene 128:135–140.

Pawson, T., and G. D. Gish. 1992. SH2 and SH3 domains: from structure to function. Cell 71:359–362.

Phizicky, E. M., R. C. Schwartz, and J. Abelson. 1986. *Saccharomyces cerevisiae* tRNA ligase. J. Biol. Chem. 261:2978–2986.

Pinkney, M., and J. G. Hoggett. 1988. Binding of the cyclic AMP receptor protein of *Escherichia coli* to RNA polymerase. Biochem. J. 250:897–902.

Porpaczy, Z., B. Sumegi, and L. Alkonyi. 1983. Association between the α-ketoglutarate dehydrogenase complex and succinate thiokinase. Biochim. Biophys. Acta 749:172–179.

Prasad, K. V. S., O. Janssen, R. Kapeller, M. Raab, L. C. Cantley, and C. E. Rudd. 1993. Src-homology 3 domain of protein kinase $p59^{fyn}$ mediates binding to phosphatidylinositol 3-kinase in T cells. Proc. Natl. Acad. Sci. USA 90:7366–7370.

Prasad, K. V. S., R. Kapeller, O. Janssen, H. Repke, J. S. Duke-Cohan, L. C. Cantley, and C. E. Rudd. 1993. Phosphatidylinositol (PI) 3-kinase and PI 4-kinase binding to the CD4-p56lck SH3 domain binds to PI 3-kinase but not PI 4-kinase. Mol. Cell Biol. 13:7708–7717.

Prelich, G., C.-K. Tan, M. Kostura, M. B. Mathews, A. G. So, K. M. Downey, and B. Stillman. 1989. Functional identity of proliferating cell nuclear antigen and a DNA polymerase-δ auxiliary protein. Nature (London) 326:517–520.

Rabidran, S. K., R. L. Haroun, J. Clos, J. Wisniewski, and C. Wu. 1993. Regulation of heat shock factor trimer formation: role of a conserved leucine zipper. Science 259:230–234.

Raff, E. C., and M. T. Fuller. 1984. Genetic analysis of microtubule function in *Drosophila*, p. 293–304. *In* G. G. Borisy, D. W. Cleveland, and D. B. Murphy (ed.), Molecular biology of the cytoskeleton. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Ratner, D. 1974. The interaction of bacterial and phage proteins with immobilized *Escherichia coli* RNA polymerase. J. Mol. Biol. 88:373–383.

Rauhut, R., P. R. Green, and J. Abelson. 1990. Yeast tRNA-splicing endonuclease is a heterotrimeric enzyme. J. Biol. Chem. 265:18180–18184.

Ray, S. K., M. Arroyo, S. Bigchi, and P. Raychaudhuri. 1992. Identification of a 60-kilodalton Rb-binding protein, RMP60, that allows the Rb-E2F complex to bind DNA. Mol. Cell. Biol. 12:4327–4333.

Redl, B., J. Walleczek, M. Stoffler-Meilicke, and G. Stoffler. 1989. Immunoblotting analysis of protein-protein crosslinks within the 50S ribosomal subunit of *Escherichia coli*: a study using dimethylsuberimidate as crosslinking reagent. Eur. J. Biochem. 181:351–356.

Ren, R., B. J. Mayer, P. Cicchetti, and D. Baltimore. 1993. Identification of a ten-amino acid proline-rich SH3 binding site. Science 259:1157–1161.

Rine, J. 1991. Gene overexpression in studies of *Saccharomyces cerevisiae*. Methods Enzymol. 194:239–251.

Rivas, G., K. C. Ingham, and A. P. Minton. 1992. Calcium-linked self-association of human complement Cls. Biochemistry 31:11707–11712.

Rivas, G., K. C. Ingham, and A. P. Minton. 1994. Ca^{2+}-linked association of human complement Cls and Clr. Biochemistry 33:2341–2348.

Rivas, G., and A. P. Minton. 1993. New developments in the study of biomolecular associations via sedimentation equilibrium. Trends Biochem. Sci. 18:284–287.

Roberts, B. L., W. Markland, A. C. Ley, R. B. Kent, D. W. White, S. K. Gutterman, and R. C. Ladner. 1992. Directed evolution of a protein: selection of potent neutrophil elastase inhibitors displayed on M13 fusion phage. Proc. Natl. Acad. Sci. USA 89:2429–2433.

Rozakis-Adcock, M., R. Fernley, J. Wade, T. Pawson, and D. Bowtell. 1993. The SH2 and SH3 domains of mammalian Grb2 couple the EGF receptor to the Ras activator mSos1. Nature (London) 363:83–85.

Russell, R. B., J. Breed, and G. J. Barton. 1992. Conservation analysis and structure prediction of the SH2 family of phosphotyrosine binding domains. FEBS Lett. 304:15–20.

Sadowski, I., J. C. Stone, and T. Pawson. 1986. A noncatalytic domain conserved among cytoplasmic protein-tyrosine kinases modifies the kinase function and transforming activity of Fujinami sarcoma virus $P130^{gag\text{-}fps}$. Mol. Cell. Biol. 6:4396–4408.

Salminen, A., and P. J. Novick. 1987. A *ras*-like protein is required for a post-Golgi event in yeast secretion. Cell 49:527–538.

Sanders, S. L., K. M. Whitfield, J. P. Vogel, M. D. Rose, and R. W. Schekman. 1992. Sec61p and BiP directly facilitate polypeptide translocation into the ER. Cell 69:353–365.

Sarkar, F. H., and S. L. Gupta. 1984. Receptors for human γ interferon: binding and crosslinking of ^{125}I-labeled recombinant human γ interferon to receptors on WISH cells. Proc. Natl. Acad. Sci. USA 81:5160–5164.

Sauer, R. T. 1979. Molecular characterization of the repressor and its gene cI. Ph.D. thesis. Harvard University, Cambridge, Mass.

Sawyers, C. L., W. Callahan, and O. N. Witte. 1992. Dominant negative MYC blocks transformation by ABL oncogenes. Cell 70:901–910.

Scherer, P. E., U. C. Manning-Krieg, P. Jeno, G. Schatz, and M. Horst. 1992. Identification of a 45-kDa protein at the protein import site of the yeast mitochondrial inner membrane. Proc. Natl. Acad. Sci. USA 89:11930–11934.

Schmidt-Dorr, T., P. Oertel-Buchheit, C. Pernelle, L. Bracco, M. Schnarr, and M. Granger-Schnarr. 1991. Construction, purification and character-

BEST AVAILABLE COPY

ization of a hybrid protein comprising the DNA binding domain of the LexA repressor and the jun leucine zipper: a circular dichroism and mutagenesis study. Biochemistry 30:9657–9664.
197. Schuster, S. C., R. V. Swanson, L. A. Alex, R. B. Bourret, and M. I. Simon. 1993. Assembly and function of a quarternary signal transduction complex monitored by surface plasmon resonance. Nature (London) 365:343–346.
198. Scott, J. K., and G. P. Smith. 1990. Searching for peptide ligands with an epitope library. Science 249:386–390.
199. Shapiro, R., and B. L. Vallee. 1991. Interaction of human placental ribonuclease with placental ribonuclease inhibitor. Biochemistry 30:2246–2255.
200. Sikela, J. M., and W. E. Hahn. 1987. Screening an expression library with a ligand probe: isolation and sequence of a cDNA corresponding to a brain calmodulin-binding protein. Proc. Natl. Acad. Sci. USA 84:3038–3042.
201. Skolnik, E. Y., B. Margolis, M. Mohammadi, E. Lowenstein, R. Fischer, A. Drepps, A. Ullrich, and J. Schlessinger. 1991. Cloning of PI3 kinase-associated p85 utilizing a novel method for expression/cloning of target proteins for receptor tyrosine kinases. Cell 65:83–90.
202. Smith, D. B., and K. S. Johnson. 1988. Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione *S*-transferase. Gene 67:31–40.
203. Smith, G. P. 1985. Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. Science 228:1315–1317.
204. Sollner, T., J. Rassow, M. Wiedmann, J. Schlossmann, P. Keil, W. Neupert, and N. Pfanner. 1992. Mapping of the protein import machinery in the mitochondrial outer membrane by crosslinking of translocation intermediates. Nature (London) 355:84–87.
205. Songyang, Z., S. E. Shoelson, M. Chaudhuri, G. Gish, T. Pawson, W. G. Haser, F. King, T. Robertys, S. Ratnofsky, R. J. Lechleider, B. G. Neel, R. B. Birge, J. E. Fajardo, M. M. Chou, H. Hanafusa, B. Schaffhausen, and L. C. Cantley. 1993. SH2 domains recognize specific phosphopeptide sequences. Cell 72:767–778.
206. Sopta, M., R. W. Carthew, and J. Greenblatt. 1985. Isolation of three proteins that bind to mammalian RNA polymerase II. J. Biol. Chem. 260:10353–10360.
207. Soutar, A. K., and D. P. Wade. 1990. Ligand blotting, p. 55–76. *In* T. Creighton (ed.), Protein function: a practical approach. IRL Press, Oxford.
208. Srere, P. A. 1987. Complexes of sequential metabolic enzymes. Annu. Rev. Biochem. 56:89–124.
209. Steitz, T. A., W. F. Anderson, R. J. Fletterick, and C. M. Anderson. 1977. High resolution crystal structures of yeast hexokinase complexes with substrates, activators, and inhibitors. J. Biol. Chem. 252:4494–4500.
210. Stephen, C. W., and D. P. Lane. 1992. Mutant conformation of p53: precise epitope mapping using a filamentous phage epitope library. J. Mol. Biol. 225:577–583.
211. Stevens, F. J. 1986. Analysis of protein-protein interaction by simulation of small-zone size-exclusive chromatography: application to an antibody-antigen association. Biochemistry 25:981–993.
212. Stoffler, G., B. Redl, J. Walleczek, and M. Stoffler-Meilicke. 1988. Identification of protein-protein cross-links within the *Escherichia coli* ribosome by immunobloting techniques. Methods Enzymol. 164:64–76.
213. Susskind, M. M., and P. Youderian. 1983. Bacteriophage P22 antirepressor and its control, p. 347–363. *In* R. W. Hendrix, J. W. Roberts, F. W. Stahl, and R. A. Weisberg (ed.), Lambda II. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
214. Tompa, P., J. Batke, J. Ovadi, G. R. Welch, and P. A. Srere. 1987. Quantitation of the interaction between citrate synthase and malate dehydrogenase. J. Biol. Chem. 282:6089–6092.
215. Traut, R. R., C. Casiano, and N. Zecherle. 1989. Crosslinking of protein subunits and ligands by the introduction of disulphide bonds, p. 101–133. *In* T. E. Creighton (ed.), Protein function: a practical approach. IRL Press, Oxford.
216. Truant, R., H. Xiao, C. J. Ingles, and J. Greenblatt. 1993. Direct interaction between the transcriptional activation domain of human p53 and the TATA box-binding protein. J. Biol. Chem. 268:2284–2287.
217. Uchiumi, T., A. J. Wahba, and R. R. Traut. 1987. Topography and stoichiometry of acidic proteins in large ribosomal subunits from *Artemia salina* as determined by crosslinking. Proc. Natl. Acad. Sci. USA 84:5580–5584.
218. Urena, J., et al. 1993. Unpublished data.
219. Vallee, R. B., and C. A. Collins. 1986. Purification of microtubules and microtubule-associated proteins from sea urchin eggs and cultured mammalian cells using Taxol, and use of exogenous taxol-stabilized brain microtubules for purifying microtubule-associated proteins. Methods Enzymol. 134:116–121.
220. Van Aelst, L., M. Barr, S. Marcus, A. Polverino, and M. Wigler. 1993. Complex formation between RAS and RAF and other protein kinases. Proc. Natl. Acad. Sci. USA 90:6213–6217.
221. Vincent, J.-P., and M. Lazdunski. 1972. Trypsin-pancreatic trypsin inhibitor association. Dynamics of the interaction and role of disulfide bridges. Biochemistry 11:2967–2977.
221a. Vinson, C. R., T. Hai, and S. M. Boyd. 1993. Dimerization specificity of the leucine zipper-containing bZIP motif on DNA binding: prediction and rational design. Genes Dev. 7:1047–1058.
222. Vinson, C. R., K. L. LaMarco, P. F. Johnson, W. H. Landschulz, and S. L. McKnight. 1988. In situ detection of sequence-specific DNA binding activity specified by a recombinant bacteriophage. Genes Dev. 2:807–806.
223. Vojtek, A. B., S. M. Hollenberg, and J. A. Cooper. 1993. Mammalian Ras interacts directly with the serine/threonine kinase Raf. Cell 74:205–214.
224. Waksman, G., D. Kominos, S. C. Robertson, N. Pant, D. Baltimore, R. B. Birge, D. Cowburn, H. Hanafusa, B. J. Mayer, M. Overduin, M. D. Resh, C. B. Rios, L. Silverman, and J. Kuriyan. 1992. Crystal structure of the phosphotyrosine recognition domain SH2 of V-src complexed with tyrosine-phosphorylated peptides. Nature (London) 358:646–653.
225. Waksman, G., S. E. Shoelson, N. Pant, D. Cowburn, and J. Kuriyan. 1993. Binding of a high affinity phosphotyrosyl peptide to the Src SH2 domain: crystal structures of the complexed and peptide-free forms. Cell 72:779–790.
226. Waldron, C., and F. Lacroute. 1975. Effect of growth rate on the amounts of ribosomal and transfer ribonucleic acids in yeast. J. Bacteriol. 122:855–865.
227. Warne, P. H., P. R. Viciana, and J. Downward. 1993. Direct interaction of Ras and the amino-terminal region of Raf-1 *in vitro*. Nature (London) 364:352–355.
228. Weber, J., R. S.-F. Lee, S. Wilke-Mounts, E. Grell, and A. E. Senior. 1993. Combined application of site-directed mutagenesis, 2-azido-ATP labeling, and *lin*-benzo-ATP binding to study the noncatalytic sites of *Escherichia coli* F_1-ATPase. J. Biol. Chem. 268:6241–6247.
229. Weiel, J., and J. W. B. Hershey. 1981. Fluorescence polarization studies of the interaction of *Escherichia coli* protein synthesis initiation factor 3 with 30S ribosomal subunits. Biochemistry 20:5859–5865.
230. Weiel, J., and J. W. B. Hershey. 1982. The binding of fluorescein-labeled protein synthesis initiation factor 2 to *Escherichia coli* 30 S ribosomal subunits determined by fluorescence polarization. J. Biol. Chem. 257:1215–1220.
231. Wells, J. A., and H. B. Lowman. 1992. Rapid evolution of peptide and protein binding properties *in vitro*. Curr. Opin. Struct. Biol. 2:355–362.
232. Weng, Z., J. A. Taylor, C. E. Turner, J. S. Brugge, and C. Seidel-Dugan. 1993. Detection of Src homology 3-binding proteins, including paxillin, in normal and v-Src-transformed Balb/c 3T3 cells. J. Biol. Chem. 268:14956–14963.
233. Wensel, T. G., and L. Stryer. 1986. Reciprocal control of retinal rod cyclic GMP phosphodiesterase by its γ subunit and transducin. Proteins Struct. Funct. Genet. 1:90–99.
234. Wensel, T. G., and S. Stryer. 1990. Activation mechanism of retinal rod cyclic GMP phosphodiesterase probed by fluorescein-labeled inhibitory subunit. Biochemistry 29:2155–2161.
235. White, R. J., and S. P. Jackson. 1992. The TATA-binding protein: a central role in transcription by RNA polymerases I, II and III. Trends Genet. 8:284–288.
236. Whyte, P., K. J. Buchkovich, J. M. Horowitz, S. H. Friend, M. Raybuck, R. A. Weinberg, and E. Harlow. 1988. Association between an oncogene and an anti-oncogene: the adenovirus E1A proteins bind to the retinoblastoma gene product. Nature (London) 334:124–129.
237. Whyte, P., N. M. Williamson, and E. Harlow. 1989. Cellular targets for transformation by the adenovirus E1A proteins. Cell 56:67–75.
238. Wiedmann, M., D. Goerlich, E. Hartmann, T. V. Kurzchalia, and T. A. Rapoport. 1989. Photocrosslinking demonstrates proximity of a 34 kDa membrane protein to different portions of preprolactin during translocation through the endoplasmic reticulum. FEBS Lett. 257:263–268.
239. Wiedmann, M., T. V. Kurzchalia, E. Hartmann, and T. A. Rapoport. 1987. A signal sequence receptor in the endoplasmic reticulum membrane. Nature (London) 328:830–833.
240. Wiman, K. G. 1993. The retinoblastoma gene: role in cell cycle control and cell differentiation. FASEB J. 7:841–845.
241. Yanofsky, C., and M. Rochmeier. 1958. The exclusion of free indole as an intermediate in the biosynthesis of tryptophan in *Neurospora crassa*. Biochim. Biophys. Acta 28:640–645.
242. Yee, S.-P., and P. E. Branton. 1985. Detection of cellular proteins associated with human adenovirus type 5 early region 1A polypeptides. Virology 147:142–153.
243. Yong, H., G. A. Thomas, and W. L. Peticolas. 1993. Metabolite-modulated complex formation between α-glycerophosphate dehydrogenase and lactate dehydrogenase. Biochemistry 32:11124–11131.
244. Young, R., and R. Davis. 1983. Yeast RNA polymerase II genes: Isolation with antibody probes. Science 222:778–782.
245. Yu, H., M. K. Rosen, T. B. Shin, C. Seidel-Dugan, J. S. Brugge, and S. L. Schreiber. 1992. Solution structure of the SH3 domain of Src and identification of its ligand-binding site. Science 258:1655–1668.
246. Zebedee, S. L., C. F. Barbas III, Y.-L. Hom, R. H. Caothien, R. Graff, J. DeGraw, J. Pyati, R. LaPolla, D. R. Burton, R. A. Lerner, and G. B. Thornton. 1992. Human combinatorial antibody libraries to hepatitis B surface antigens. Proc. Natl. Acad. Sci. USA 89:3175–3179.
247. Zervos, A. S., J. Gyuris, and R. Brent. 1993. Mxi1, a protein that specifically

BEST AVAILABLE COPY

interacts with Max to bind Myc-Max recognition sites. Cell 72:223–232.
Zhang, X.-F., J. Settleman, J. M. Kyriakis, E. Takeuchi-Suzuki, S. J. Elledge, M. S. Marshall, J. T. Bruder, U. R. Rapp, and J. Avruch. 1993. Normal and oncogenic p21ras proteins bind to the amino-terminal regulatory domain of c-Raf-1. Nature (London) 364:308–313.
Zhou, M., S. Felder, M. Rubinstein, D. R. Hurwitz, A. Ullrich, I. Lax, and J. Schlessinger. 1993. Real-time measurements of kinetics of EGF binding to soluble EGF receptor monomers and dimers support the dimerization model for receptor activation. Biochemistry 32:8193–8198.
250. Zimmerman, J. K., and G. K. Ackers. 1971. Molecular sieve studies of interacting protein systems: behavior of small zone profiles for reversibly self-associating solutes. J. Biol. Chem. 246:7289–7292.

BEST AVAILABLE COPY

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :
C12Q 1/68, C12P 21/02
G01N 33/53, 33/531, A61K 35/14

A1

(11) International Publication Number: WO 91/04339

(43) International Publication Date: 4 April 1991 (04.04.91)

(21) International Application Number: PCT/US90/05155

(22) International Filing Date: 12 September 1990 (12.09.90)

(30) Priority data:
408,767 18 September 1989 (18.09.89) US

(71) Applicant: CALIFORNIA BIOTECHNOLOGY INC. [US/US]; 2450 Bayshore Parkway, Mountain View, CA 94043 (US).

(72) Inventors: CORDELL, Barbara ; 4051 Ben Lomond Drive, Palo Alto, CA 94306 (US). WOLF, David ; 2142 Bellview Drive, Palo Alto, CA 94303 (US).

(74) Agents: MURASHIGE, Kate, H. et al.; Irell & Manella, 545 Middlefield Road, Suite 200, Menlo Park, CA 94025 (US).

(81) Designated States: AT (European patent), AU, BE (European patent), CA, CH (European patent), DE (European patent)*, DK (European patent), ES (European patent), FR (European patent), GB (European patent), IT (European patent), JP, LU (European patent), NL (European patent), SE (European patent).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: ASSAYS AND REAGENTS FOR AMYLOID DEPOSITION

(57) Abstract

The present invention provides an in vitro tissue culture-based assay for amyloid deposition specific for Alzheimer's disease which is suitable for routine drug screening analysis. Immunological diagnostic reagents for Alzheimer's disease are also provided.

* See back of page

DESIGNATIONS OF "DE"

Until further notice, any designation of "DE" in any international application whose international filing date is prior to October 3, 1990, shall have effect in the territory of the Federal Republic of Germany with the exception of the territory of the former German Democratic Republic.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	ES	Spain	MC	Monaco
AU	Australia	FI	Finland	MG	Madagascar
BB	Barbados	FR	France	ML	Mali
BE	Belgium	GA	Gabon	MR	Mauritania
BF	Burkina Faso	GB	United Kingdom	MW	Malawi
BG	Bulgaria	GR	Greece	NL	Netherlands
BJ	Benin	HU	Hungary	NO	Norway
BR	Brazil	IT	Italy	PL	Poland
CA	Canada	JP	Japan	RO	Romania
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo			SE	Sweden
CH	Switzerland	KR	Republic of Korea	SN	Senegal
CM	Cameroon	LI	Liechtenstein	SU	Soviet Union
DE	Germany	LK	Sri Lanka	TD	Chad
DK	Denmark	LU	Luxembourg	TG	Togo
				US	United States of America

ASSAYS AND REAGENTS FOR AMYLOID DEPOSITION

Field of the Invention

The present invention relates to assays and reagents useful for the chemical intervention of amyloidosis in Alzheimer's disease.

Background of the Invention

Alzheimer's disease (AD) is an age-related brain degenerative disease that is the most common cause of intellectual failure in late life. Neuritic or senile plaques and neurofibrillary tangles (NFT) are the hallmark characteristic of the histopathology of Alzheimer's brains. These plaques and tangles are believed to result from deposits of two different proteins which share the properties of the amyloid class of proteins specific for AD.

The major protein component of amyloid is an ~4 kilodalton (kd) protein, designated the beta-protein or A4 protein due to a partial beta pleated structure or its molecular weight, respectively. The amino acid sequence of A4 has been defined (Wong et al., (1985) Proc Natl Acad Sci USA 82:8729-8732) and full-length cDNA encoding a primary translation product of 695 residues has been cloned (Kang et al., (1987) Nature 325:733-736) while other cDNAs have been identified which encode a 751-residue or 770-residue precursor form (Ponte et al., (1988) Nature 331:525-527; Tanzi et al., (1988) Nature

331:528-530; and Kitaguchi et al., (1988) Nature 331:530-532).

The A4 protein accumulates extracellularly, both in brain parenchyma and in the walls of blood vessels, generally as amyloid plaques which form aggregate fibril structures and are insoluble on SDS-polyacrylamide gels. The fibrils are generally identified as amyloid based on their green birefringence after staining with Congo red and their 40- to 90-A diameter.

The second protein, mentioned previously, accumulates intracellularly in neurons of Alzheimer's brains (Castano and Frangione, (1988) Lab Invest 58:122-132) and forms tangles composed of structures resembling paired helical filaments (PHFs). In contrast to the beta-amyloid protein, the primary structure and number of proteins comprising PHFs are unknown. PHF-containing neurites are found in the periphery of the plaque, whereas deposits of beta-amyloid protein form the central core of mature plaques, surrounded by degenerated neurites and glial cells.

Although the etiology of AD is unknown, it has been demonstrated that the frequency of neuritic plaques found in the cortex of AD patients correlates with the degree of dementia (Roth et al., (1966) Nature 209:109-110; Wilcock and Esiri, (1982) J Neurol Sci 56:343-356). The therapeutic goals in amyloidosis are to prevent further deposition of amyloid material and to promote or accelerate its resorption. To date, there are no means available to treat the pathogenesis of AD and the paucity of understanding concerning the mechanism of amyloid formation in AD is a major obstacle in the development and design of therapeutic agents that can intervene in this process. Moreover, no animal models for brain amyloidosis with beta-amyloid protein deposits or PHFs exist, creating yet another obstacle to test such putative therapeutic agents.

Logical therapeutic approaches are now, however, emerging for treating the particular amyloidosis associated with AD. These approaches are attributable, in part, from the successes and failure gained in attempting to treat other forms of amyloidosis, such as the use of dimethyl sulfoxide which blocks amyloid formation from Bence Jones proteins in vitro (Coria et al., (1988) Lab Invest 58:454-458) and use of colchicine to reduce the size of renal amyloid deposits and induce clinical remissions in several cases of familial Mediterranean fever and amyloid nephropathy (Ravid et al., (1977) Ann Intern Med 87:568-570).

Efforts directed to the design of in vitro models of age-related cerebral amyloidogenesis using A4-derived synthetic peptides are disclosed in Castano et al., (1986) Biochem Biophys Res Comm 141:782-789, and in Kirschner et al., (1987) Proc Natl Acad Sci USA 84:6953-6957. Castano et al. demonstrated that amyloid fibrils could be formed in vitro when using a synthetic peptide corresponding to the amino-terminal 28 residues of the amyloid core protein. This 28 residue peptide, as well as a 17 residue sequence contained within the 28 amino acids, both formed fibrils which stain similarly to material isolated from Alzheimer's brains; however, the synthetic amyloid fibrils were soluble, unlike the naturally occurring insoluble amyloid isolated from Alzheimer's brains. Kirschner et al. demonstrated that the same 28 residue peptide could be produced as an insoluble aggregate; however, this particular in vitro model is not expected to correlate well to the cellular environment in which amyloid deposition occurs.

Dyrks et al., (1988) EMBO J 7:949-957 showed that a shortened cell-free translation product comprising the amyloid A42 part and the cytoplasmic domain of the 695-residue precursor can form multimers. While aggregation was observed employing an in vitro cell-free

system, this system fails to reveal whether aggregation of the translation product would naturally follow in vivo. Moreover, the in vitro cell-free system does not address protein stability issues, that is, whether adequate levels of the protein could be expressed, whether protein proteolysis exists, and other concerns generally associated with in vivo expression of recombinant proteins.

Therefore, there exists a need for a definitive cellular deposition model with which one may assay agents capable of chemically intervening in the process of amyloid deposition. Such a method should be relatively simple to perform and should be highly specific in distinguishing AD plaques from the plaques associated with other disorders. Furthermore, it is desirable that the assay be capable of being reduced to a standardized format. The present invention satisfies such needs and provides further advantages.

Summary of the Invention

The present invention provides a method for determining the ability of a potential therapeutic agent to intervene in the amyloid deposition process associated with Alzheimer's disease in a cellular environment, which method utilizes a recombinantly produced amyloid substrate in a screening assay. The present invention also allows for the development and use of immunological reagents to detect the formation of preamyloid protein aggregation in the cell lines provided by the invention.

To achieve the objects and in accordance with the purpose of the invention, as embodied and broadly described herein, a method of screening agents capable of intervention in Alzheimer's disease amyloidosis comprises:

a) culturing a cell line capable of expressing a gene encoding beta-amyloid protein under conditions suit-

able to produce the beta-amyloid protein as an insoluble, preamyloid aggregate;

b) combining a known quantity of the agent to be tested to the cell culture; and

c) monitoring the combination to determine whether preamyloid aggregate formation is reduced.

In an alternative embodiment of the invention, preamyloid formation can be induced through infection of a cell line with a recombinant virus capable of expressing the beta-amyloid protein as an insoluble preamyloid aggregate. Such recombinant viruses carry expression vectors comprising DNA encoding the beta-amyloid protein.

Immunoassay kits employing the reagents useful to screen potential amyloid intervening agents are also provided by the present invention.

Brief Description of the Drawings

FIG. 1 is a schematic illustration of two amyloid expression constructs employing the vaccinia pUV1 insertion vector.

FIG 2. illustrates the results of immunoprecipitation of ^{35}S-methionine labeled VV:A99 infected CV-1 cell lysates using APCP antibodies. The arrows mark A99 protein.

FIG 3. are fluorescent photomicrographs of infected CV-1 cells stained with APCP antibodies. FIG. 3A is a Mock control; FIG. 3B is a VV:CONT control; FIG. 3C is the VV:99 construct; and FIG. 3D is the VV:42 construct. The magnification is 200x with a 0.4 second exposure time for each photo.

FIG. 4 is a illustration of the modified beta-actin expression selection vector, pAX-neo, that was employed to express the beta-amyloid core constructs in mammalian cells.

Detailed Description of the Preferred Embodiments

As indicated above, the invention involves a method of screening agents capable of intervention in Alzheimer's disease amyloidosis.

As used herein, the term "beta-amyloid core protein" or "A4 protein" refers to an approximately 4 kd protein first identified by Glenner and Wong, (1984) Biochem Biophys Res Comm 120:885, which is defined at the amino terminus by sequence analysis as a mixture of four peptides with slightly different amino termini, the amino termini of the three smaller peptides being completely encoded by that of the largest.

The term "beta-amyloid precursor protein" refers to either the amyloid precursor protein of 695 amino acids (Kang et al., (1987) supra) or the 751 amino acid protein (Ponte et al., (1988) supra) containing within their sequence, the beta-amyloid core protein sequence defined above. The A4 core protein begins at amino acid 597 of the 695 amino acid protein and at amino acid 653 of the 751 amino acid sequence.

The terms "preamyloid aggregation", "preamyloid formation", or "preamyloid deposits" refer to a morphological description -- first discovered by Tagliavini et al., (1988) Neurosci Lett 93:191-196 -- of spherical, granular deposits which are considerably smaller than pre-plaques and plaques found at a high frequency in the brains of Alzheimer's victims. These deposits can be occasionally detected with silver stain but not with Congo red, a stain to which amyloid proteins demonstrate high binding affinity.

As used herein, the term "insertion vector" includes plasmids, cosmids or phages capable of mediating homologous recombination into a viral genome such that the DNA encoding the beta-amyloid protein is stably carried by the resulting recombinant virus. In one embodiment of the

invention plasmids constructed from vaccinia virus DNA are employed.

The term "expression vector" includes plasmids, cosmids or phages capable of synthesizing a protein encoded by the respective recombinant gene carried by said vector. Such vectors are independently replicated in or capable of integration into the chromosome of an appropriate host cell for expression of the amyloid protein.

A cell has been "transformed" by exogenous or heterologous DNA when such DNA has been introduced inside the cell. The transforming DNA may or may not be integrated (covalently linked) into chromosomal DNA making up the genome of the cell. In prokaryotes, yeast, and mammalian cells, for example, the transforming DNA may be maintained on an episomal element such as a plasmid. The cell has been stably transformed when the cell is able to establish cell lines or clones comprised of a population of daughter cells containing the transforming DNA. A "clone" is a population of cells derived from a single cell or common ancestor by mitosis. A "cell line" is a clone of a cell that is capable of stable growth in vitro for many generations.

A. Beta-Amyloid Coding Sequences

The beta-amyloid genes may be synthetic or natural, or combinations thereof. The gene encoding the natural 751 amino acid precursor protein is described in PCT WO88/03951, published 2 June 1988 and assigned to the same assignee of the present application, and the expression of the protein in mammalian cells is provided in Example 4 therein. The relevant portions of this publication are specifically incorporated herein by reference.

The genes encode the A42 core protein or an amyloid protein, A99, which comprises the A42 core protein and the cytoplasmic domain. This latter protein consists of the 42 residue core protein and 57 amino acids of the

cytoplasmic domain of the 751 precursor protein. The sequence of A99 is as follows:

Asp Ala Glu Phe Arg His Asp Ser Gly Tyr (10) Glu Val His His Gln
Lys Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys Gly Ala
Ile Ile Gly Leu Met Val Gly Gly Val Val Ile Ala Thr Val Ile
Val Ile Thr Leu Val Met Leu Lys Lys Lys Gln Tyr Thr Ser Ile
His His Gly Val Val Glu Val Asp Ala Ala Val Thr Pro Glu Glu
Arg His Leu Ser Lys Met Gln Gln Asn Gly Tyr Glu Asn Pro Thr
Tyr Lys Phe Phe Glu Gln Met Gln Asn.

These genes are provided for expression of the desired protein using recombinant DNA expression vectors.

As mentioned above, these genes may be natural, synthetic or combinations thereof. When preparing a synthetic nucleotide sequence, it may be desirable to modify the natural amyloid nucleic acid sequence. For example, it will often be preferred to use codons which are preferentially recognized by the desired host. In some instances, it may be desirable to further alter the nucleotide sequence, either synthetic or natural, to create or remove restriction sites to, for example, enhance insertion of the gene sequence into convenient expression vectors or to substitute one or more amino acids in the resulting polypeptide to increase stability. A general method for site-specific mutagenesis is described in Noren et al., (1989) Science 244:182-188.

Peptides of this precursor protein, for example, those derived from the A4 core protein, are also provided herein for the generation of specific immunological re-

agents and may also be synthetic or natural. Synthetic oligonucleotides are prepared by either the phosphotriester method as described by Edge et al., (1981) Nature 292:756 and Duckworth et al., (1981) Nuc Acids Res 9:1691 or the phosphoramidite method as described by Beaucage and Caruthers, (1981) Tet Lett 22:1859 and Matteucci and Caruthers, (1981) J Am Chem Soc 103:3185, and can be prepared using commercially available automated oligonucleotide synthesizers.

B. Vaccinia Viral Vectors

The coding sequences for the amyloid proteins can be inserted into vaccinia virus plasmid insertion vectors for the purpose of generating recombinant vaccinia viruses using the methods described in Moss et al., (1983) Methods in Gene Amplification, Vol. 3, Elsevier-North Holland, p. 202-213; and in Moss et al., (1984) J Virol 49:857-864. The amyloid-vaccinia recombinants can then be used for (1) expression of the respective amyloid protein and analysis of preamyloid formation, and (2) production of amyloid antibodies.

The two vaccinia virus insertion vectors, pSC11 (Chakrabarti et al., (1985) Mol Cell Biol 5:3403-3409 and pUV1 (Falkner et al., (1987) Nuc Acids Res 15:7192) were used for the expression of the amyloid proteins and generation of amyloid-vaccinia recombinants. Both vectors are of the co-insertion variety and each contains two vaccinia virus promoters. One promoter (P1) is used to drive the expression of a selectable marker gene (in this case, beta-galactosidase) while the other promoter (P2) is used to drive expression of the heterologous amyloid DNA insert. Both are flanked by vaccinia virus DNA (an interrupted thymidine kinase [tk] gene) which facilitates homologous recombination into a wild-type vaccinia virus genome and provides a selection mechanism (generation of tk minus viruses). The pSC11 vector utilizes a vaccinia

early-late promoter (P7.5) to drive heterologous gene expression and has a single SmaI cloning site. The pUV1 vector utilizes a vaccinia late promoter (P11) to drive heterologous gene expression and is designed for the expression of fusion proteins behind the ATG of the P11 late gene. In all cases, amyloid-pUV1 constructs were made using the most 5' (after the ATG) cloning site (EcoRI) in order to avoid introduction of additional amino terminal amino acids into the native amyloid protein sequence.

C. Recombinant Expression Vectors and Hosts

It will also be understood by those skilled in the art that both procaryotic and eucaryotic systems may be used to express the amyloid genes described herein. Procaryotes most frequently are represented by various strains of E. coli; however, other microbial strains may also be used. Plasmid vectors which contain replication sites, selectable markers and control sequences derived from a species compatible with the host are used; for example, E. coli is typically transformed using derivatives of pBR322, a plasmid derived from an E. coli species by Bolivar et al., (1977) Gene 2:95. pBR322 contains genes for ampicillin and tetracycline resistance, and thus provides multiple selectable markers which can be either retained or destroyed in constructing the desired vector. Commonly used procaryotic control sequences which are defined herein to include promoters for transcription initiation, optionally with an operator, along with ribosome binding site sequences, include such commonly used promoters as the beta-lactamase (penicillinase) and lactose (lac) promoter systems (Chang et al., (1977) Nature 198:1056), the tryptophan (trp) promoter system (Goeddel et al., (1980) Nucleic Acids Res 8:4057), the lambda-derived P_L promoter (Shimatake et al., (1981) Nature 292:128) and N-gene ribosome binding site, and the

trp-lac (trc) promoter system (Amann and Brosius, (1985) Gene 40:183).

In addition to bacteria, eucaryotic microbes, such as yeast, may also be used as hosts. Laboratory strains of Saccharomyces cerevisiae, Baker's yeast, are most used although a number of other strains or species are commonly available. Vectors employing, for example, the 2 micron origin of replication of Broach, (1983) Meth Enz 101:307, or other yeast compatible origins of replication (see, for example, Stinchcomb et al., (1979) Nature 282:39; Tschumper et al., (1980) Gene 10:157 and Clarke et al., (1983) Meth Enz 101:300) may be used. Control sequences for yeast vectors include promoters for the synthesis of glycolytic enzymes (Hess et al., (1968) J Adv Enzyme Reg 7:149; Holland et al., (1978) Biochemistry 17:4900). Additional promoters known in the art include the promoter for 3-phosphoglycerate kinase (Hitzeman et al., (1980) J Biol Chem 255:2073). Other promoters, which have the additional advantage of transcription controlled by growth conditions and/or genetic background are the promoter regions for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, the alpha factor system and enzymes responsible for maltose and galactose utilization. It is also believed terminator sequences are desirable at the 3' end of the coding sequences. Such terminators are found in the 3' untranslated region following the coding sequences in yeast-derived genes.

It is also, of course, possible to express genes encoding polypeptides in eucaryotic host cell cultures derived from multicellular organisms. See, for example, Axel et al., U.S. Patent No. 4,399,216. These systems have the additional advantage of the ability to splice out introns and thus can be used directly to express genomic fragments. Useful host cell lines include VERO, HeLa, baby hamster kidney (BHK), CV-1, COS, MDCK, NIH 3T3, L,

and Chinese hamster ovary (CHO) cells. Expression vectors for such cells ordinarily include promoters and control sequences compatible with mammalian cells such as, for example, the commonly used early and late promoters from Simian Virus 40 (SV40) (Fiers et al., (1978) Nature 273:113), or other viral promoters such as those derived from polyoma, herpes virus, Adenovirus 2, bovine papilloma virus, or avian sarcoma viruses. The controllable promoter, hMTII (Karin et al., (1987) Nature 299:797-802) may also be used. General aspects of mammalian cell host system transformations have been described by Axel, supra.

Insect expression systems may also be employed to express the amyloid genes. For example, the baculovirus polyhedrin gene has been employed for high-level expression of heterologous proteins (Smith et al., (1983) Mol Cell Biol 3(12):2156-2165; Summers et al., "Genetic Engineering of the Genome of the Autographa Californica Nuclear Polyhedrosis Virus", Banbury Report: Genetically Altered Viruses in the Environment, 22:319-339, Cold Spring Harbor Laboratory, 1985).

D. Generation of Stably Transfected Cell Lines

The amyloid DNA clones expressed in vaccinia can also be used to generate stably transfected cell lines expressing the amyloid proteins. In general, these cell lines are generated by first constructing one of two expression plasmids. In both expression plasmids, the selectable marker is provided by a G418 neomycin expression cassette (neo) consisting of the SV40 early promoter, the bacterial kanamycin-resistance gene also containing its own promoter, the SV40 intervening sequence, and the SV40 polyadenylation site from the early region. In the first expression plasmid, the amyloid DNA cloning site is flanked at the 5' end by the human metallothionein gene promoter, pMtIIa, modified with an SV40 enhancer, and at the 3' end by the SV40 polyadenylation site from the early

region. In the second expression construct, the amyloid DNA cloning site is flanked at the 5' end by a beta-actin promoter, and at the 3' end by a sequence encoding a useful polyadenylation site, such as that of the SV40 early region or the beta-actin gene.

Each of the vectors described above can be transformed into a mammalian cell line such as, but not limited to, those described in the following examples by either calcium phosphate-DNA coprecipitation or electroporation. A day later, the cells are subjected to 1 mg/ml G418 to provide pools of G418-resistant colonies. Successful transformants, also having a stable inheritance of the DNA contained in the expression construct, are then plated at low density for purification of clonal isolates. Clonal isolates are then analyzed for maximum production of the amyloid protein of interest and high-producing clones are expanded to serve as stock.

E. Detection Methods for Preamyloid Formation

The diagnosis of amyloidosis is established by demonstration of the characteristic emerald-green birefringence of tissue specimens stained with Congo red and examined by polarization microscopy. Congo red staining is generally carried out using commercially available diagnostic kits. The isolation and characterization of the A4 protein has allowed specific antibodies to be raised that recognized cerebral amyloid in Alzheimer's disease (Allsop et al (1986) Neurosci Lett 68:252-256). Moreover, Tagliavini et al., (1988) supra, have demonstrated that antibodies can be generated which detect in both Alzheimer's patients and to a lesser extent in non-demented individuals preamyloid deposits, which deposits lack the tinctorial and optical properties of amyloid and are, therefore, undetectable using conventional staining methods employing principally Congo red, but also thioflavin S or silver salts.

Standard protocols can be employed for preparing antibodies directed against the amyloid proteins of the invention. Techniques for preparing both polyclonal and monoclonal antibodies are well known in the art. Briefly, polyclonal antibodies are prepared by injecting amyloid protein or synthetic amyloid peptides with an adjuvant into an animal such as rabbits or mice. The amyloid protein may need to be conjugated to a carrier protein such as bovine serum albumin or keyhole limpet hemacyanin using a chemical process which employs carbodiimide, glutaraldehyde, or other cross-linking agents. Alternatively, the protein may be administered without being conjugated to a carrier protein. Vaccinia virus recombinants which are expressing amyloid proteins may also be used to prepare antibodies. The vaccinia virus recombinants are injected into an animal and then the animal is boosted several weeks after the initial immunization. Ten days to two weeks later the animals are bled and antiserum is collected and analyzed for titer.

Monoclonal antibodies are commonly prepared by fusing, under appropriate conditions, B-lymphocytes of an animal which is making polyclonal antibodies with an immortalizing myeloma cell line. The B-lymphocytes can be spleen cells or peripheral blood lymphocytes. Techniques for fusion are also well known in the art, and in general, involve mixing the cells with a fusing agent such as polyethylene glycol. Successful hybridoma formation is assessed and selected by standard procedures such as, for example, HAT medium. From among successful hybridomas, those secreting the desired antibody are screened by assaying the culture medium for their presence.

Standard immunological techniques such as ELISA (enzyme-linked immunoassay), RIA (radioimmunoassay), IFA (immunofluorescence assay) and Western blot analysis, which are well known in the art, can be employed for diagnostic screening for amyloid expression. A vast

literature now exists with respect to various modifications of the basic assay principle, which is simply that there must be a specific association between target analyte and antibody, which association is detectable qualitatively and/or quantitatively. Fluorescent, enzymatic, or radioactive labels are generally used.

One typical arrangement utilizes competition, between labeled antigen (e.g. amyloid protein) and the analyte, for the antibody, followed by physical separation of bound and unbound fractions. Analyte competes for the binding of the labeled antigen; hence more label will remain in the unbound fraction when larger amounts of analyte are present. In this competitive-binding type assay, the sample is incubated with a known titer of labeled amyloid protein and amyloid protein antibody. Antibody-protein complex is then separated from uncomplexed reagents using known techniques and the amount of label in the complexed material is measured, e.g. by gamma counting in the case of radioimmunoassay or photometrically in the case of enzyme immunoassay. The amount of amyloid protein in the sample, if any, is determined by comparing the measured amount of label with a standard curve.

Other embodiments of this basic principle include use of labeled antibodies per se, sandwich assays involving a three-way complex between analyte, anti-analyte antibody, and anti-antibody wherein one of the components contains a label, and separation of bound and unbound fractions using an immunosorbent. Agglutination assays which result in visible precipitates are also available (Limet et al., (1982) J Clin Chem Clin Biochem 20:142-147).

F. Screening Assay

The present assay provides one of the first steps in addressing the question whether preamyloid corti-

cal deposits herald organic dementia. The concomitant appearance of preamyloid deposits and senile plaques suggests that preamyloid deposits may evolve into senile plaques.

Down's syndrome is the one known disease closely related to the proposed etiology of AD. As from their twenties onward, Down's patients develop the full spectrum of Alzheimer's changes, i.e., NFTs, congophilic angiopathy and senile plaques. As reported in Giacione et al., (1989) Neurosci Letts 97:232-238, a time-related analysis of preamyloid deposits and senile plaque distribution showed an age-dependent, inverse correlation between extracellular preamyloid deposits and senile plaque in Down's patients. While a similar, time-dependent study with Alzheimer's patients cannot be conducted, it is expected that a corresponding pattern (preamyloid turning to senile plaque deposits) would be found. Therapeutic agents which interfere with this process promise the development of successful therapeutic regimens for Alzheimer's disease.

In the practice of the method of the invention, the expression of the amyloid protein is initiated by culturing the transformed cell line under conditions which are suitable for cell growth and expression of the amyloid protein. In this method, high level expression of the protein is preferred. In one embodiment of the invention, a CHO cell line transformed with a beta-actin vector comprising the DNA encoding the A42 or A99 amyloid protein is grown in a mammalian culture medium such as, for example, a 1:1 mixture of F12 medium and DME medium with 10% fetal calf serum for 5-72 hr at 37°C. Transfected viral monolayers are selected and plaque purified, and stocks of amyloid-vaccinia recombinant viruses are prepared.

2° ref to virtues of present

The formation of the preamyloid aggregates can be monitored by standard immunocytochemical methods using,

for example, beta-amyloid primary antibodies which are detected using a secondary, labeled anti-antibody. If one is interested in testing whether the compound of interest can inhibit preamyloid formation, the compound is introduced to the culture medium before monitoring for preamyloid aggregation. Alternatively, the compound is introduced to the culture medium after preamyloid formation has been established and this reaction mixture is monitored to see whether the compound induces amyloid resorption.

Potential therapeutic compounds for use in the present invention include, for example, amyloid-fibril denaturing agents such as dimethyl sulfoxide, and cytotoxic agents such as colchicine and chlorambucil. The efficacy of these agents may be monitored through observation of reduced antibody binding to the amyloid deposit. Reduction in such binding is indicative of reduced preamyloid deposition. Alternatively, preamyloid formation in the host cell may trigger other cellular events which could be employed as markers unrelated to the etiology of Alzheimer's disease, but correlative with the presence of preamyloid deposits. For example, an increase in the level of certain enzymes, specifically proteases, may be measured in lieu of the preamyloid deposition. Typically, an increase in the concentration levels of these enzymes is observed when cultured cells are subjected to stress.

The present invention also encompasses kits suitable for the above diagnostic or screening methods. These kits contain the appropriate reagents and are constructed by packaging the appropriate materials, including the preamyloid protein aggregates immobilized on a solid support with labeled antibodies in suitable containers, along with any other reagents (e.g., wash solutions, enzyme substrate, anti-amyloid antibodies) or other materials required for the conduct of the assay.

The reagents are usually premeasured for ease of use. An optional component of the kit is a set of instructions describing any of the available immunoassay methods. For example, a kit for a direct assay can comprise preamyloid proteins aggregates immobilized on a solid immunoassay support and a container comprising labeled antibody to the amyloid protein, as well as the other reagents mentioned above.

The following examples are designed to elucidate the teachings of the present invention, and in no way limit the scope of the invention. Most of the techniques which are used to transform cells, construct vectors perform immunoassays, and the like are widely practiced in the art, and most practitioners are familiar with the standard resource materials which describe specific conditions and procedures. The examples are written in observation of such knowledge and incorporate by reference procedures considered conventional in the art.

EXAMPLE 1

Description of Amyloid Plaque Core DNA Constructs

The following examples describe the expression vectors containing the 42 amino acid plaque core region (A42), and the 42 amino acid plaque core region including the 57 amino acid adjacent carboxy-terminal region of the beta-amyloid precursor protein (A99). Alternative constructs for the A42 and A99 constructs were prepared which included a 17 amino acid amyloid signal sequence. As these constructs did not express the amyloid protein well, further experimentation with these vectors was not performed.

Recombinant vaccinia viruses bearing amyloid DNAs encoding each of the two amyloid constructs (VV:A42 and VV:A99) were generated by standard methods as reviewed by Mackett and Smith in (1986) J Gen Virol 67:2067-2082,

which is incorporated herein by reference. FIG. 1 illustrates the various amyloid expression constructs, all of which were modified at the 5' end to satisfy the cloning constraints of the vaccinia P11 promoter in the pUV1 vector. Specifics for each construct are as follows:

A. VV:A42:

The A42-encoding sequence (nucleotides 2080 to 2205, numbered in accordance with the 751 amyloid precursor sequence) was synthesized as a 145 basepair (bp) EcoRI-BamHI oligomer, provided below, containing the appropriate TGA stop codon and an amino-terminal Asn-Ser adaptor sequence:

5' AAT TCC GAT GCA GAA TTC CGA CAT GAC TCA
GGA TAT GAA GTT CAT CAT CAA AAA TTG GTG
TTC TTT GCA GAA GAT GTG GGT TCA AAC AAA
GGT GCA ATC ATT GGA CTC ATC GTG GGC GGT
GTT GTC ATA GCG TGA TCT AGA TGA G 3'

The synthetic fragment was ligated to EcoRI- and BamHI-digested pGem1 (Promega-Biotec), deriving pGemA42. The EcoRI-BamHI fragment of pGemA42 was subsequently isolated and ligated into the EcoRI-BamHI site of pUV1 deriving pUV1:A42.

The XbaI-SalI fragment of pUV1:A42 (287bp) was further subcloned into mp18 for sequence confirmation.

B. VV:A99:

The DNA encoding the amyloid protein for the pUV1-A99 constructs was derived from 4T4B, a plasmid encoding the 751 amino acid precursor protein. The construction of plasmid 4T4B is described in Example 3 of PCT/US87/02953, owned by the same assignee. The relevant portions of this publication are incorporated herein by reference. The 590 bp DdeI-PvuII fragment of plasmid 4T4B

was isolated from the carboxy-terminal 1 kilobase (kb) EcoRI fragment of 4T4B and ligated with a 27 bp EcoRI-DdeI adaptor sequence and cloned into the EcoRI- and SmaI-digested pUV1, deriving pUV1:A99.

The 761 bp XbaI-SalI fragment of pUV1:A99 was further subcloned into the XbaI-SalI vector fragment of mp18 and pGem2. Sequence data confirmed the predicted sequence.

EXAMPLE 2

Expression of Amyloid Proteins

The vaccinia insertion vectors described in Example 1 were used to generate amyloid-vaccinia recombinant viruses as follows.

A. Preparation of Amyloid-Vaccinia Virus Recombinants

Confluent monolayers of CV-1 cells in 60 mm dishes were infected with vaccinia virus (Wyeth strain) at a multiplicity of infection (moi) of 0.05 pfu/cell. At 0.5 hr post-infection, the cells were transfected with a calcium phosphate precipitate of 10 ug insertion plasmid DNA and 0.5 ug wild-type vaccinia virus DNA. Cells were fed with complete medium and incubated at 37°C for two days. Monolayers were collected and TK- vaccinia viruses were selected on TK-143 cells in the presence of 5-bromodeoxyuridine (BudR) at 25 ug/ml. At 48 hr after infection, monolayers were overlaid with 1% agarose containing 300 ug/ml 5-bromo-4-chloro-3-indolyl-B-D-galactopyranoside (Xgal). At 4-6 hr, blue plaques were picked and further purified by two additional rounds of plaque purification in the presence of BudR and Xgal. Stocks of the amyloid-vaccinia recombinant viruses were prepared in TK-142 or CV-1 cells. Recombinant viral DNA was prepared from each stock and was shown by Southern blot analysis to contain the appropriate amyloid DNA

insert and to be free of contamination with wild-type or spontaneous TK- vaccinia.

B. Identification of Amyloid-specific Polypeptides Produced By Vaccinia Virus Recombinants

Characterization of the CV-1 expressed VV:A42 and VV:A99 amyloid proteins was carried out employing immunoprecipitation and polyacrylamide gel analysis of ^{35}S-methionine-labeled infected cell protein using antibodies directed against the carboxy-terminal region of the amyloid precursor.

The beta-amyloid antibodies were generated from synthetic peptides. The synthetic peptides were prepared using solid phase synthesis according to standard protocols. Purification of the crude peptides was accomplished by desalting with gel filtration followed by ion-exchange chromatography and preparative reverse-phase liquid chromatography. Each peptide was fully characterized by amino acid composition and sequence analysis. COOH-CORE corresponds to amino acids 653-680(DAEFRHDSGYEVHHQKLVFFAEDVGSSA) (the carboxy-terminal two amino acids were taken from the amino acid sequence of Masters et al., (1985) Proc Natl Acad Sci 82:4245-4249 and are different in the deduced translation of the A4 cDNA of Ponte et al., supra. COOH-B2 and COOH-C2 correspond to amino acids 736-751(NGYENPTYKFFEQMQN), COOH-B3 and COOH-C3 correspond to amino acids 705-719(KKKQYTSIHHGVVEV) and COOH-C5 corresponds to amino acids 729-742(HLSKMQQNGYENPT). Reference for the numbering of peptides along the topology of the A4 precursor is from Ponte et al., supra. New Zealand white rabbits were immunized intradermally with 500 ug of peptide conjugated to keyhole limpet hemocyanin. The rabbits were first bled at 4 weeks and 1 week later the rabbits were boosted with 250 ug conjugated peptide. Subsequent bleeds were done at 3 week intervals with boosts following 1 week later. All

animals were treated in accordance with institutional guidelines. Antibody titers against the appropriate peptide were determined by enzyme-linked immunosorbent assays coupled with horseradish peroxidase and found to be $7.4x10^4$, $2.7x10^5$, $1x10^5$, $9.1x10^6$, $8.2x10^5$, and $2.5x10^5$ for COOH-CORE, COOH-B2, COOH-C2, COOH-B3, COOH-C3, and COOH-C5, respectively.

Antibodies to 9523 correspond to amino acids 673-685(AEDVGSKNGAIIG) and 9524 correspond to amino acids 701-712(LVMLKKKQYTSI). Antibodies to these two peptides were generated by coinjecting New Zealand white rabbits each with 200 ug methylated bovine serum albumin (PBS) plus 200 ug of the respective synthetic peptide in PBS. Rabbits were boosted one, two and three weeks following primary inoculation with identical amounts of peptide. Serum samples were taken at week 6 and titered against APCP synthetic peptide. Titers achieved were $1.5x10^4$ for 9523 and $4x10^5$ for 9524.

CV-1 cells were infected with VV:99 at a multiplicity of infection of one. ^{35}S-methionine (250 uCi/ml) was added at 20 hr post infection for 4 hr. Cell lysates were prepared and aliquots containing 10^7 cpm were immunoprecipitated with amyloid-specific antisera (COOH-B3, COOH-C5 and COOH-CORE) or normal rabbit serum and protein A.

Immunoprecipitates of ^{35}S-methionine cell lysates were analyzed on denaturing 20% SDS-polyacrylamide gels. As shown in FIG. 2, high levels of expression and stability of the A99 protein generated by VV:A99 was demonstrated. The control sera (normal, nonimmune rabbit sera) did not display reactivity with the VV:A99 protein product. The VV:A99 amyloid core protein migrated as a broad band spanning approximately 11.5-17 kd molecular weight. In addition, higher molecular weight forms of the A99 protein were clearly observed.

The expression product of VV:A99 demonstrated high level expression of the 99 amino acid core protein and showed evidence of self-aggregation as well as aggregation with other proteins or self-aggregation combined with proteolysis since multimers of A99 did not always occur in integers of 11.5-17 kd.

EXAMPLE 3

Staining of A42 and A99 Expressing Cells

Two human, SK-N-MC (ATCC # HTB10) and IMR-32, (ATCC # CCL127) and one rat, PC-12 (Green and Tischler, (1976) Proc Natl Acad Sci USA 73:2424-2428) neuronal cell lines were examined for their ability to permit efficient infection with the VV:A42 and VV:A99 recombinant viruses. All cell lines were documented as permissive hosts for vaccinia virus replication by infecting cells with a given amount (moi=2) of vaccinia virus of known titer. The infected cells were harvested 20 hours after infection, disrupted by freeze-thaw, and then titered. The yield was compared to the input viral units and if 20-100 fold increase results, the host cell was considered permissible for vaccinia replication.

These neuronal lines and the CV-1 cell line were employed for amyloid staining studies. The culture medium for each host was as follows:

CV-1: The medium was Eagle MEM supplemented with 10% FBS, penicillin, streptomycin and L-Gln.

SK-N-MC: Eagle MEM supplemented with 10% FBS, non-essential amino acids, penicillin, streptomycin and L-Gln.

PC-12: DMEM21, 5% DHS, 5% DFBS and L-Gln; and

IMR-32: Eagle MEM (Hank's BSS) and 10% deltaFBS plus nonessential amino acids, penicillin, streptomycin and L-Gln.

Each cell line was grown to confluency on a microscope slide divided into 4 individual chambers (Lab Tech). One chamber was mock infected, the second infected with a control recombinant virus lacking A4 sequences (VV:CONT), the third chamber infected with VV:A99, and the fourth chamber infected with VV:A42. This is an internally controlled method since each slide was manipulated as a single unit.

Viral infections were carried out at a moi from 5 to 20 viral plaque forming units (pfu) per cell and were harvested for staining at approximately 20 hours post infection. Slides prepared for immunocytochemistry were fixed with 4% paraformaldehyde and permeabilized with 0.2% Triton X-100 prior to treatment with primary and rhodamine-conjugated second antibodies (Capell Labs). Briefly, after permeabilization, cells were washed with PBS containing 0.2% gelatin. 100 ul of primary amyloid antibody (diluted 1/200 with PBS plus 0.2% gelatin) was incubated on the cells at 37^{o}C for 30 minutes. Cells were washed for 10 minutes in PBS and 0.2% gelatin, then incubated at 37^{o}C for 20 minutes with a 1/200 dilution (in PBS and gelatin) of secondary antibody (goat-anti-rabbit) tagged with Rhodamine. Cells were washed for 10 min in PBS and gelatin, then mounted for visualization in a fluorescent microscope. Antibodies used with success included 9523, 9524, B3 and C5. CORE antibodies were not assessed. Alternatively, the slides were fixed in 4% paraformaldehyde then stained with Thioflavin S or Congo red, and counterstained with hematoxylin according to directions in commercial kits (Sigma).

IMR-32 and PC-12 cells presented some technical difficulties and thus further investigation with these cell lines was terminated. The IMR-32 cells did not adhere well to the microscope slides, which could be alleviated by pretreatment with laminin, and, moreover, the IMR-32 cells did not tolerate the serum-free conditions during the infections. PC-12 cells showed high background immunostaining, hence, differences between experimental and control samples were not dramatic.

FIG. 3 shows fluorescent photomicrographs of CV-1 cells stained with 1/200 dilutions of the core domain antibodies 9523 antibodies. Specific and robust staining was seen in only the VV:A99 and VV:42 infected cells. VV:99 specific staining, but not VV:42 staining, was seen with the B3 antibody as would be anticipated since this region is not included in the VV:A42 construct (results not shown). Faint punctate staining was observed for both antibodies on all cells presumably due to endogenous A4 precursor expression. The VV:A99 and VV:A42 infected cells displayed strong reactivity in the form of large deposit-like structures which are cell associated. The deposit-like structures are probably not cell debris from the viral cytopathicity since they are not seen in the VV:CONT cells and their immunoreactivity could be eliminated by preadsorption of the antisera with the synthetic peptide used to raise the serum.

The possible potentiating effect of aluminum on deposit formation was investigated by pretreating the cells with 50 mM $AlCl_3$. Aluminum might be considered a "cofactor" in the pathology of amyloid formation since it is present in plaques. However, no obvious qualitative difference in the degree of deposit formation between cultures treated and untreated with aluminum was found.

It seems relevant that several researchers investigating A4 core domain immunoreactivity in brains of Alzheimer's victims describe similar structures as those

in FIG. 3. Each group reported finding significant amounts of specifically stained spherical, granular deposits which were considerably smaller than pre-plaques and plaques (Davies et al., (1988) Neurolog 38:1688-1693; Ikeda et al., (1989) Lab Invest 60:113-122; Tagliavini et al., (1988) supra; Tate-Ostroff et al., (1989) Proc Natl Acad Sci 86:745-749). All research groups independently propose that the observed small granular deposits are the very early stages of amyloid plaque development. The structures observed in our cell culture system are analogous to those seen in the Alzheimer's diseased brain. It was noted by these investigators that the granular deposits could be occasionally detected with silver stain but not with Congo red. Because the Alzheimer's granular deposits were highly reactive with A4 antisera but were not easily reacted with stains capable of recognizing the tinctorial properties of amyloid, the structures were termed "preamyloid" deposits.

EXAMPLE 4

Establishment of Stable Cell Lines

A number of constructs expressing the beta-amyloid core protein were constructed using a derivative of the beta-actin expression/selection vector designated pHbetaAPr-1-neo. This vector, illustrated in FIG. 4, is a combination of the following elements:

a) bp 1-4300 is the 4.3 kb EcoRI-AluI fragment from the human beta-actin gene isolate p14Tbeta-17 (Leavitt et al., (1984) Mol Cell Biol 4:1961-1969). For sequencing details of the promoter see Ng et al., (1985) Mol Cell Biol 5:2720-2732. The cap site, 5' untranslated region and IVS 1 positions are indicated in FIG. 4. There is no ATG codon present in the 5' UT nor in the polylinker region from the 3' splice site to the BamHI site;

b) bp 4300-4320 is in part derived from pSP64 polylinker (Melton et al., (1984) Nuc Acids Res 12:7035-7056);

c) bp 4320-6600 is derived from pcDV1 (Okayama & Berg, (1983) Mol Cell Biol 3:280-289); and

d) bp 6600-10000 is the PvuII-EcoRI fragment from pSV-neo (Southern & Berg, (1982) J Mol App Genet 1:327-341) containing the bacterial neomycin gene linked to the SV40 origin plus early promoter. The direction of transcription is as indicated in FIG. 4. This vector was altered by deleting the EcoRI site and adding a new EcoRI site within the polylinker 3' to the SalI site and 5' to the HindIII site. This modified vector is designated pAX-neo. Beta-actin A42 was constructed by excising the EcoRI-BamHI 145 bp fragment from pGEM-A42, adding a SalI-EcoRI adaptor sequence (5'-TCG ACA TGG ATG CAC AAT TA-3') and cloning into the pAX-neo expression vector at the SalI-and BamHI sites. The beta-actin A99 plasmid was constructed by excising the 670 bp EcoRI-HindIII fragment of $pGEM_2$-A99, adding the above-described SalI-EcoRI adaptor sequence and cloning into the pAX-neo vector at the SalI and HindIII sites.

Each construct was introduced into CHO cells by the calcium phosphate precipitation method using 7 ug of each DNA per 10^6 cells, and a resistant population was selected with G418-neomycin. The efficiency of transfection for the A99 or A42 constructs was over 10^3 for 10^6 cells and pools of cells transfected with either beta-actin A99 or with beta-actin A42 were selected using G418-neomycin resistance (500 ug/ml).

Cell lysates from these pools are prepared and analyzed by immunoprecipitation of the A4 proteins as well as by Western blotting. High expressing clones are then selected and assayed for "preamyloid" deposits using the immunocyto-staining procedures described in Example 3.

EXAMPLE 5

Assay for Preamyloid Deposition

Cells infected with VV:99 or VV:42 which are capable of forming amyloid deposits are plated in a 96-well microtiter plate. To make the appropriate dilutions and additions, an automated pipetter is used to introduce the drug to be tested to the cells. A range of concentrations of the drug is incubated in a tissue culture incubator (or preincubated) with the cells at 37°C for a predetermined time period, or alternatively, for 3 to 72 hours.

Following incubation, the culture media is removed, and the cells are prepared for preamyloid measurement as follows. The cells are fixed for immunocytochemical staining with amyloid antibodies. The primary antibodies are introduced followed by incubation with labeled, secondary anti-antibodies and the level of binding between the primary and secondary antibodies is measured using an ELISA plate reader to record the optical density of the labeled antibody. A smaller optical density reading as compared to a control sample of cells grown in the absence of the test drug is indicative of that drug's ability to inhibit amyloid deposition. This procedure may be modified to permit detection of preamyloid dissolution using a correlative enzyme marker.

It will be apparent to those skilled in the art that various modifications and variations can be made in the method of the present invention without departing from the scope or spirit of the invention. Thus, it is intended that the claims cover the modifications and variations of the invention.

What is claimed is:

1. A method of screening agents capable of intervention in Alzheimer's disease amyloidosis comprising:

a) culturing a cell line capable of expressing a gene encoding beta-amyloid protein under conditions suitable to produce the beta-amyloid protein as an insoluble, preamyloid aggregate;

b) combining a known quantity of the agent to be tested to the cell culture; and

c) monitoring the combination to determine whether preamyloid aggregate formation is reduced.

2. The method of claim 1 wherein the beta-amyloid gene encodes a protein comprising the amyloid plaque core domain.

3. The method of claim 1 wherein the beta-amyloid gene encodes a protein comprising the amyloid plaque core and the carboxy-terminal domains.

4. The method of claim 3 wherein the beta-amyloid gene encodes the following polypeptide:

10
Asp Ala Glu Phe Arg His Asp Ser Gly Tyr Glu Val His His Gln
20 30
Lys Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys Gly Ala
40 (42)
Ile Ile Gly Leu Met Val Gly Gly Val Val Ile Ala Thr Val Ile
50 60
Val Ile Thr Leu Val Met Leu Lys Lys Lys Gln Tyr Thr Ser Ile
70
His His Gly Val Val Glu Val Asp Ala Ala Val Thr Pro Glu Glu
80 90
Arg His Leu Ser Lys Met Gln Gln Asn Gly Tyr Glu Asn Pro Thr
(99)
Tyr Lys Phe Phe Glu Gln Met Gln Asn.

5. The method of claim 1 wherein the preamyloid is derived from stable cell lines infected with recombinant vaccinia virus comprising the beta-amyloid gene.

6. The method of claim 1 wherein the cell lines are derived from mammalian host cells.

7. The method of claim 1 wherein the agent to be tested is introduced during the growth phase of the cell culture to determine whether the agent inhibits preamyloid plaque formation.

8. The method of claim 1 wherein the agent to be tested is added to the amyloid aggregate to determine whether the agent dissolves preamyloid plaque formation.

9. An immunological reagent capable of detecting preamyloid aggregate formation.

10. The immunological reagent of claim 9 which is a monoclonal antibody.

11. The immunological reagent of claim 9 which is a polyclonal antibody.

12. A kit for an immunoassay to screen compounds capable of chemical intervention in amyloidosis of Alzheimer's disease comprising:

a predetermined amount of preamyloid aggregate specific for Alzheimer's disease; and

a predetermined amount of labeled antibody to said preamyloid aggregate.

13. The kit according to claim 12 wherein said label is a component of an enzymatic reaction.

14. The kit according to claim 12 wherein said preamyloid aggregate is immobilized on a solid immunoassay support.

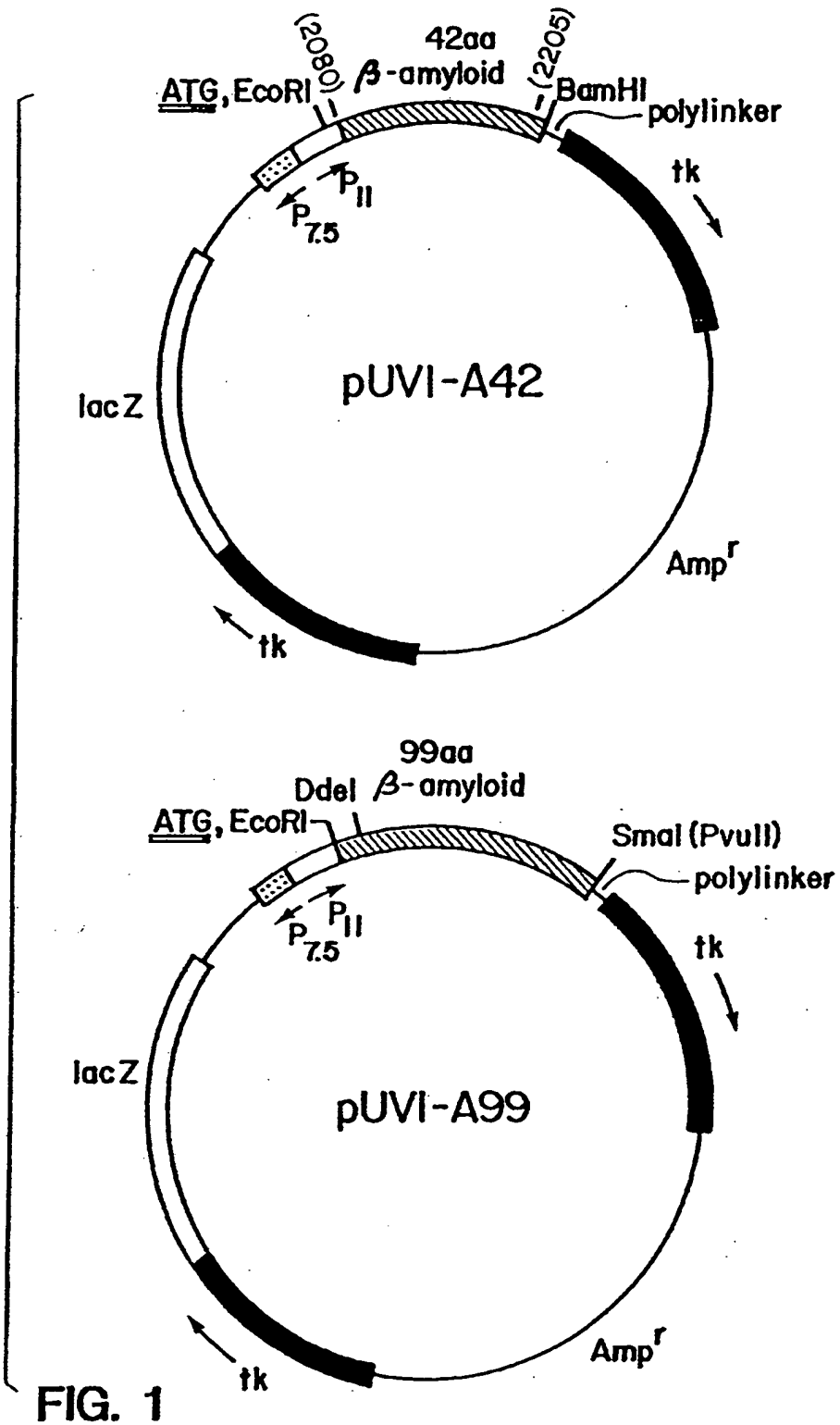

(2080)
42aa
β-amyloid
(2205)
ATG, EcoRI
BamHI
polylinker
tk
P11
P7.5
lacZ
pUVI-A42
Amp^r
tk
99aa
DdeI
β-amyloid
ATG, EcoRI
SmaI (PvuII)
polylinker
tk
P11
P7.5
lacZ
pUVI-A99
Amp^r
tk

FIG. 1

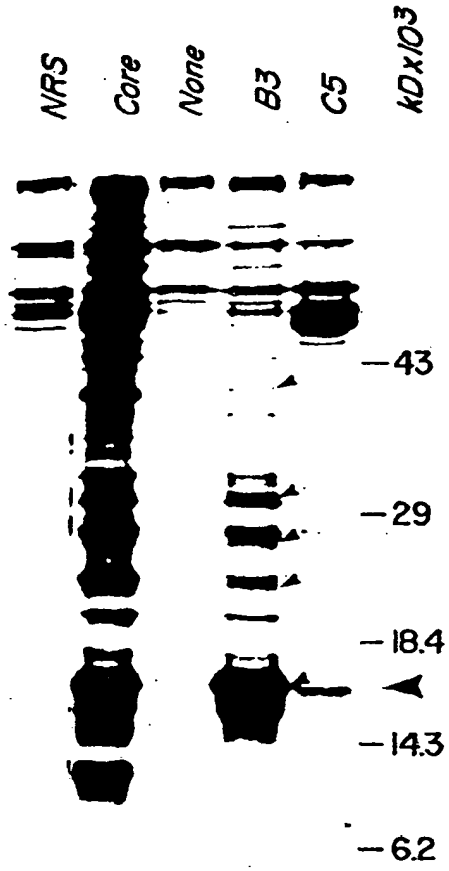


FIG.2

FIG.3A

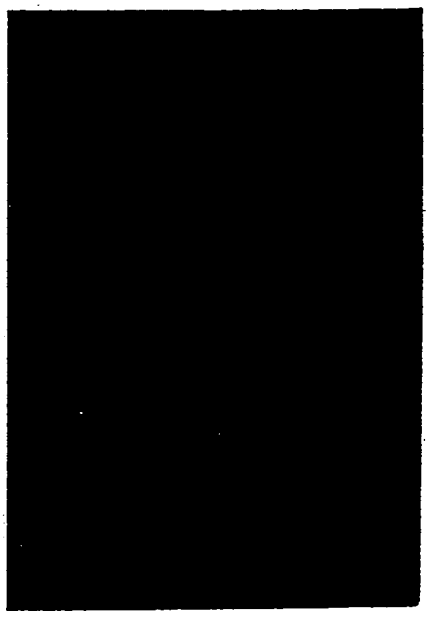


FIG.3B

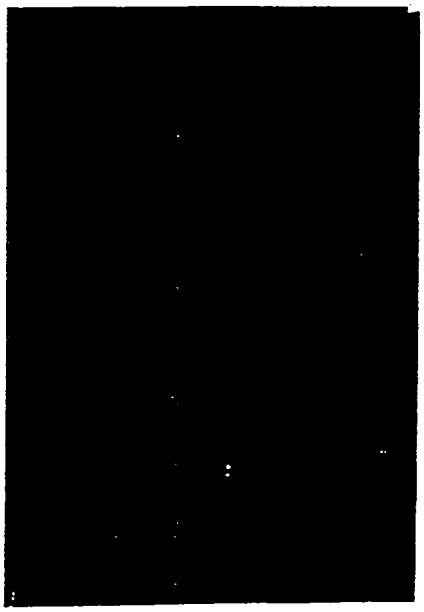


FIG.3C

FIG.3D

BEST AVAILABLE COPY

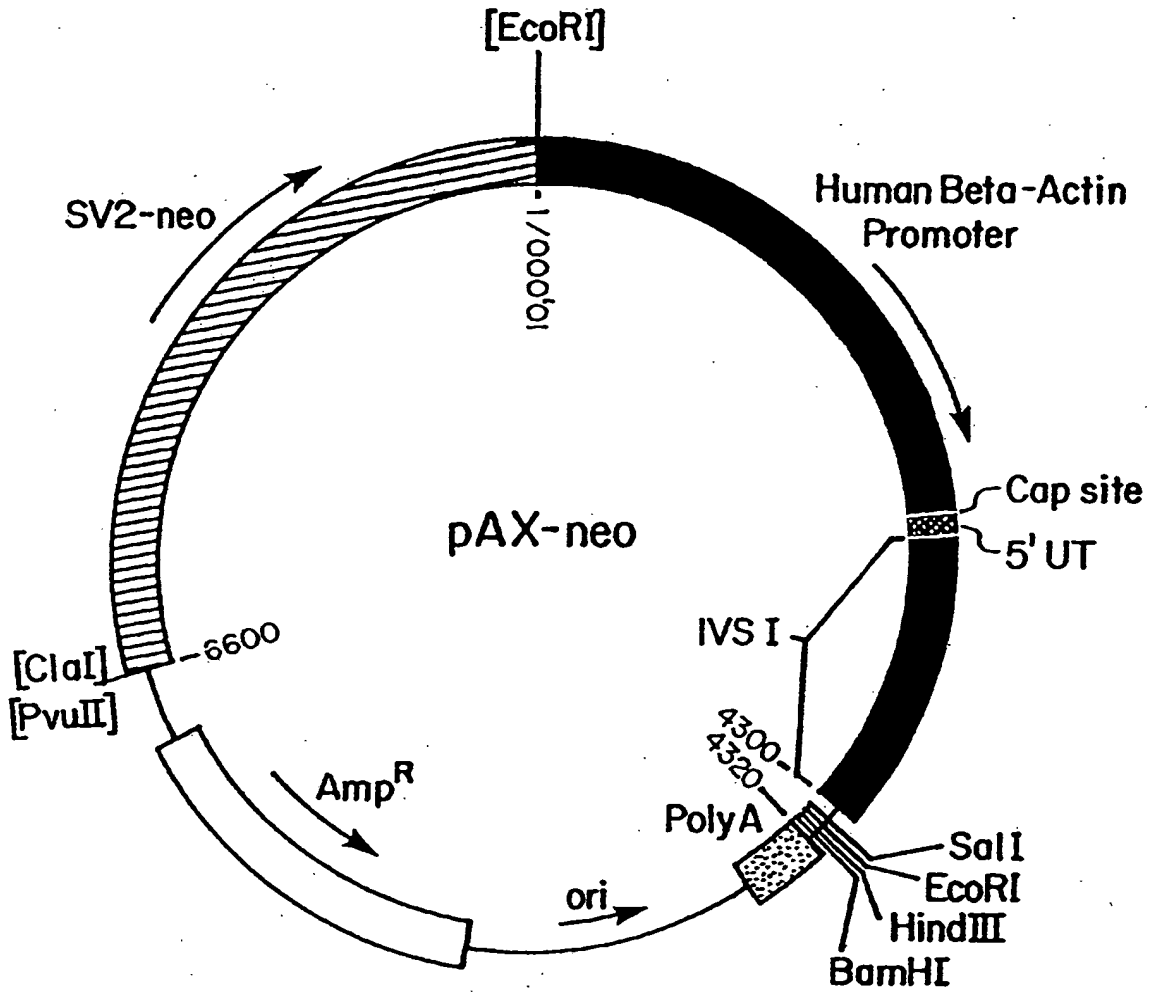


FIG. 4

BEST AVAILABLE COPY

INTERNATIONAL SEARCH REPORT

International Application No PCT/US90/05155

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) [3]

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC(5): C12Q 1/68; C12P 21/02; G01N 33/53, 33/531; A61K 35/14
US.CL: 435/6, 7, 70.1, 172.3, 810; 436/543, 548, 808; 530/350, 387, 839

II. FIELDS SEARCHED

Minimum Documentation Searched [4]

Classification System [5]	Classification Symbols
U.S. CL.	435/6, 7, 70.1, 172.3, 810; 436/543, 548, 808; 530/350,387, 839

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched [6]

APS. CAS. BIOSIS

III. DOCUMENTS CONSIDERED TO BE RELEVANT [14]

Category [15]	Citation of Document, [16] with indication, where appropriate, of the relevant passages [17]	Relevant to Claim No. [18]
A, P	US,A, 4,919,915 (AVERBACK) 24 April 1990, See column 1, lines 15-22; column 1, line 59-column 2, line 32; column 2, lines 41-46 and 53-55 and column 9, lines 25-55.	1-14
A, P	US,A, 4,912,206 (GOLDGABER ET AL.) 27 March, 1990, see Figures 1 and 3 and claims 1-3.	1-6
X Y	US,A, 4,666,829 (GLENNER ET AL.) 19 May 1987, see Abstract and column 4, lines 5-11 and 19-29.	9-11 12-14
A	US,A, 4,264,729 (BELJANSKI) 28 April 1981, See Abstract and claim 1.	1-8
Y	WO,A, 89/07657 (NEVE ET AL.) 24 August 1989, see page 15, line 15-page 16, line 7.	9-14

[*] Special categories of cited documents: [13]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search [2]	Date of Mailing of this International Search Report [2]
26 November 1990	24 JAN 1991
International Searching Authority [1]	Signature of Authorized Officer [20]
ISA/US	Janelle Graeter Janelle Graeter

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)

Category *	Citation of Document, ** with indication, where appropriate, of the relevant passages **	Relevant to Claim No **
A Y	EP,A, 0,304,013 (KITAGUCHI ET AL) 22 February 1989, see Abstract, page 5, lines 43-44; page 7, line 45-page 8, line 5; page 8, line 30-page 9, line 12; page 11, line 47-page 12, line 31 and page 16, line 1-page 17, line 8.	1-8 9-14

BEST AVAILABLE COPY

Proc. Natl. Acad. Sci. USA
Vol. 94, pp. 6618–6622, June 1997
Biochemistry

Prion-inducing domain 2–114 of yeast Sup35 protein transforms *in vitro* into amyloid-like filaments

CHIH-YEN KING*, PETER TITTMANN†, HEINZ GROSS†, ROLAND GEBERT‡, MARKUS AEBI§, AND KURT WÜTHRICH*¶

*Institut für Molekularbiologie und Biophysik, †Institut für Zellbiologie, ‡Institut für Pflanzenwissenschaften, and §Mikrobiologisches Institut, Eidgenössische Technische Hochschule, CH-8093 Zürich, Switzerland

Contributed by Kurt Wüthrich, April 14, 1997

ABSTRACT **The yeast non-Mendelian genetic factor [PSI], which enhances the efficiency of tRNA-mediated nonsense suppression in *Saccharomyces cerevisiae*, is thought to be an abnormal cellular isoform of the Sup35 protein. Genetic studies have established that the N-terminal part of the Sup35 protein is sufficient for the genesis as well as the maintenance of [PSI]. Here we demonstrate that the N-terminal polypeptide fragment consisting of residues 2–114 of Sup35p, Sup35pN, spontaneously aggregates to form thin filaments *in vitro*. The filaments show a β-sheet-type circular dichroism spectrum, exhibit increased protease resistance, and show amyloid-like optical properties. It is further shown that filament growth in freshly prepared Sup35pN solutions can be induced by seeding with a dilute suspension of preformed filaments. These results suggest that the abnormal cellular isoform of Sup35p is an amyloid-like aggregate and further indicate that seeding might be responsible for the maintenance of the [PSI] element *in vivo*.**

Self-propagating protein conformational changes have been proposed to be the cause of transmission of mammalian transmissible spongiform encephalopathies (1, 2), as well as for two yeast non-Mendelian inheritance elements, [PSI] and [URE3] (3–7). In the case of [PSI], an altered conformation of the Sup35 protein (Sup35p), which is the yeast homolog of the eukaryotic translation termination factor eRF3, is thought to be the determinant (3, 8). Consistent with this proposal it was observed that the maintenance of [PSI] only requires the N-terminal 114-amino acid domain of Sup35p, and that overproduction of Sup35p or its N-terminal fragment in yeast induces the *de novo* appearance of [PSI] (9–11). Differential sedimentation and fluorescence microscopic studies further established a correlation between Sup35p coalescence and the appearance of [PSI], suggesting that ordered aggregation converts newly synthesized Sup35p into its like and is thus responsible for the propagation of the [PSI] element (12, 13). In this study, we investigate the properties of the Sup35p polypeptide fragment consisting of residues 2–114, Sup35pN. It is found that Sup35pN aggregates to form amyloid-like filaments *in vitro*. We then show that seeding with preexisting filaments can speed up the formation of filaments from freshly prepared Sup35pN solutions.

EXPERIMENTAL PROCEDURES

Protein Purification and Filament Preparation. A DNA fragment encoding the N-terminal 114-residue segment of Sup35p, with an extra Met–Gly–Ser_2–His_6–Ser_2–Gly_2–Ser segment at the N terminus and a stop codon at the C terminus, was obtained by PCR from yeast genomic DNA, using appropriate oligonucleotides. The fragment was inserted into the expression vector pMW172 (14) and the protein was overexpressed in the BLR21(DE3)/pLysS *Escherichia coli* strain (Novagen). Cell lysate was prepared in 20 mM Tris·HCl buffer (pH 7.9) containing 0.5 M NaCl, 5 mM imidazole, and 6 M guanidine hydrochloride (GdmCl), and loaded onto a Ni^{2+}-NTA affinity column (Qiagen, Chatsworth, CA) followed by reverse-phase high performance liquid chromatography (HPLC, C18 column) in 0.1% trifluoroacetic acid (TFA) using an acetonitrile gradient. The N-terminal Met–Gly–Ser_2–His_6–Ser_2–Gly_2–Ser–Met segment was removed by cyanogen bromide cleavage (15). Unreacted peptide was separated from Sup35pN by a NaCl gradient, using an SP-Sepharose column (Pharmacia) in 20 mM sodium acetate (pH 4.5) containing 6 M urea, which was removed immediately after separation using a C18 reverse-phase HPLC column in 0.1% TFA with an acetonitrile gradient. Final buffer exchanges were achieved by gel filtration chromatography on Sephadex G25 columns (Pharmacia). The primary structure of Sup35pN was confirmed by electrospray mass spectrometry, sequencing of the N-terminal 25 residues, and isoelectric focusing. Filaments of Sup35pN were prepared in 0.1% (vol/vol) TFA/40% (vol/vol) acetonitrile using reverse-phase HPLC fractions containing isocratically eluted Sup35pN. A typical preparation of a 100 μM solution of Sup35pN yielded filaments after 1 week of incubation at 4°C. Spontaneous filament formation was also observed in 50 mM sodium phosphate buffer (pH 2.0) with 40% acetonitrile. No influence on filament formation was seen when borosilicate, polystyrol, or polypropylene tubes were used.

Electron Microscopy (EM). Samples (5 μl) of the filament suspension were adsorbed to glow discharged carbon-coated copper grids. These were washed twice with deionized water, negatively stained with 2% (wt/vol) uranyl acetate, and air-dried after removal of excess liquid. The specimens were examined in a Philips CM12 transmission electron microscope at 100 kV, and images were recorded with a Gatan 694 slow scan CCD camera.

Circular Dichroism (CD) Spectroscopy. CD spectra were recorded on a Jasco J710 spectropolarimeter at 22°C. All samples were briefly sonicated before measurement.

Polarization Light Microscopy. Filaments were sedimented at 20,000 × *g*, washed with buffer A (50 mM sodium phosphate, pH 7.0/150 mM NaCl), stained with 50 μM Congo red in buffer A at room temperature for 1 min, washed in succession with buffer A and deionized water, then placed on glass slides, dried at room temperature, and covered. The samples were studied using a Zeiss Photomicroscope III equipped with crossed polars.

Protease K Resistance Assay. About 8 μg of Sup35pN filaments were suspended in 18 μl Tris/EDTA (TE) buffer (10 mM Tris·HCl/1 mM EDTA, pH 8.0) containing 0.8 M GdmCl.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

© 1997 by The National Academy of Sciences 0027-8424/97/946618-5$2.00/0

Abbreviations: Sup35p, Sup 35 protein; Sup35pN, Sup35 protein fragment of residues 2–114; CD, circular dichroism; TFA, trifluoroacetic acid; GdmCl, guanidinium hydrochloride; EM, electron microscopy.

¶To whom reprint requests should be addressed.

Two microliters of protease K (Boehringer Mannheim no. 1413 783) solution in TE buffer at concentrations from 0.5 to 4.0 μg/ml were added and the mixture was incubated at 37°C for 80 min. The reaction was terminated by adding 5 mM phenylmethylsulfonyl fluoride. Twenty microliters of SDS gel loading buffer (100 mM Tris·HCl, pH 6.8/4% SDS/0.2% bromophenol blue/20% glycerol) were added and the samples were boiled for 3 min. The resulting solution was briefly centrifuged and then subjected to 15% tricine–SDS/PAGE, which has particularly high resolution for small polypeptides (16). Protein was detected by Coomassie brilliant blue staining.

Seeding Assay. Freshly prepared Sup35pN solutions were subjected to addition of different preparations of protein aggregates. In the case of seeding in 0.1% TFA/40% acetonitrile, a 50 μM solution of Sup35pN containing 1% (wt/wt) of protein aggregates as seeds was incubated at 4°C. In the case of seeding in 50 mM sodium phosphate at pH 7.0, a 5 μM solution of Sup35pN containing 5% of aggregates as seeds was incubated at 22°C. After overnight undisturbed incubation, aggregates were sedimented at 20,000 × *g* for 20 min, dissolved in 10 μl SDS gel loading buffer containing 6 M urea, boiled for 3 min, separated on 15% SDS polyacrylamide gels, and detected by Coomassie brilliant blue staining. For CD measurements, 2% (wt/wt) filaments were added as seeds to a freshly prepared 5 μM Sup35pN solution in 50 mM sodium phosphate at pH 7.0, and the mixture was incubated overnight at 22°C.

RESULTS

Sup35pN was overexpressed in *E. coli* and purified to homogeneity as judged by SDS/PAGE. It had limited solubility in aqueous buffers and, after sonication, submillimolar aqueous suspensions contained amorphous aggregates, as judged by EM. In 40% acetonitrile/60% H_2O at pH 2.0 the formation of amorphous aggregates could be suppressed, and after prolonged incubation unbranched filaments with a diameter of 3 nm and variable lengths were observed (Fig. 1*a*). It was also noted that two or more filaments could further associate laterally to form fibrillar structures with larger diameter (Fig. 1*b*).

Sup35pN undergoes extensive secondary structural changes upon aging. A freshly prepared solution exhibits a far UV CD spectrum that indicates little α-helix or β-sheet content. In contrast, the spectrum of an aged solution shows β-sheet-like characteristics, with a single differential absorption minimum near 220 nm (Fig. 2*a*). These β-type spectral features are predominantly associated with the filaments, since the supernatant obtained after the filaments were removed from the aged solution by centrifugation had a similar CD spectrum as the freshly prepared Sup35pN solution (Fig. 2*b*). Amorphous aggregates of Sup35pN, which show strong light scattering, do not give CD signals. Freshly prepared and aged samples were indistinguishable by SDS/PAGE or isoelectric focusing under denaturing conditions (data not shown), indicating that the covalent polypeptide structure in the freshly prepared protein solution is preserved during aging.

When Sup35pN filaments were collected by centrifugation and washed with sodium phosphate buffer at neutral pH, numerous red-stained aggregates were observed by light microscopy after Congo red treatment (Fig. 3*a*). Similar to amyloids, the stained aggregates exhibited green–yellow color when examined with crossed polars (Fig. 3*b*), showing that the filaments of Sup35pN are amyloid-like and contain regular secondary structure (17, 18).

To test whether the *in vitro* generated filaments had increased protease resistance, we subjected the following two samples to protease K digestion. (*i*) Filaments were sedimented from an aged Sup35pN solution and resuspended in Tris/EDTA buffer (pH 8.0) containing 0.8 M GdmCl. (*ii*) As a control, the same number of filaments were denatured in 8 M GdmCl and then diluted 10-fold using Tris/EDTA buffer, whereby amorphous aggregation of denatured protein was induced by carrying out the dilution very slowly. The experiments showed that Sup35pN filaments

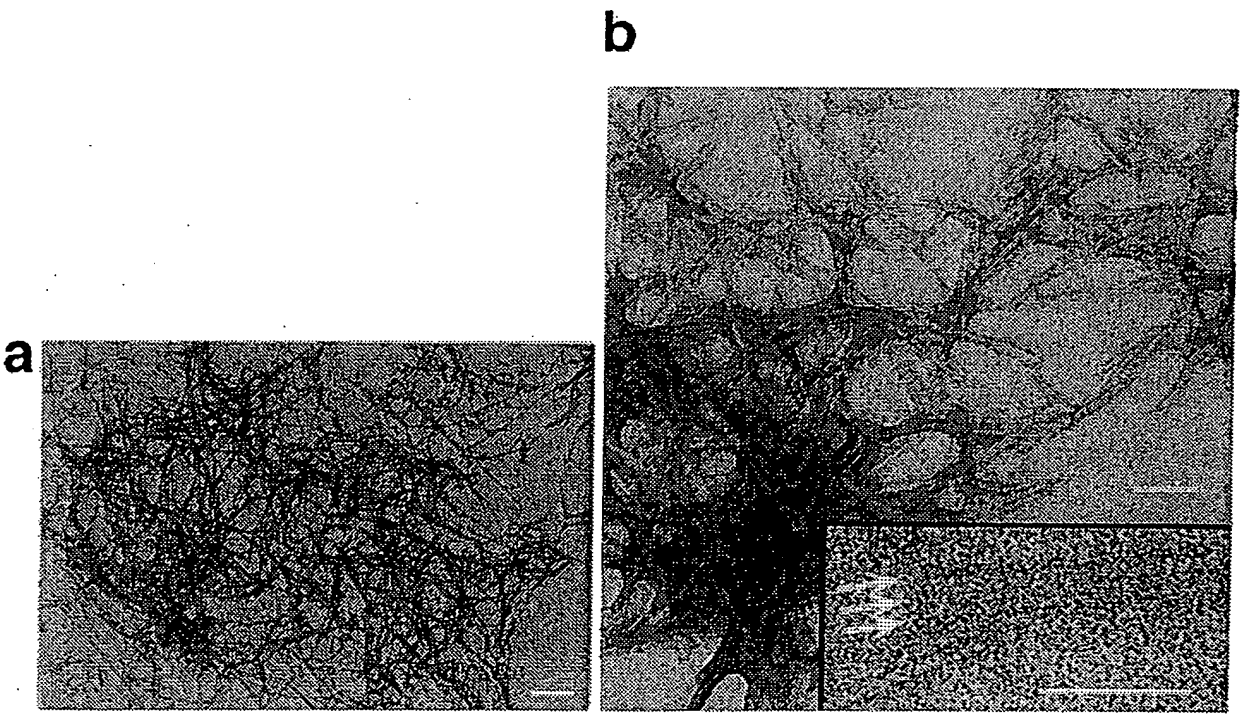


FIG. 1. (*a* and *b*) Electron micrographs of negatively stained Sup35pN filaments. (*b Inset*) Laterally associated fibrils are marked by arrows. (Bars = 250 nm in *a* and 100 nm in *b* and the *Inset*.)

BEST AVAILABLE COPY

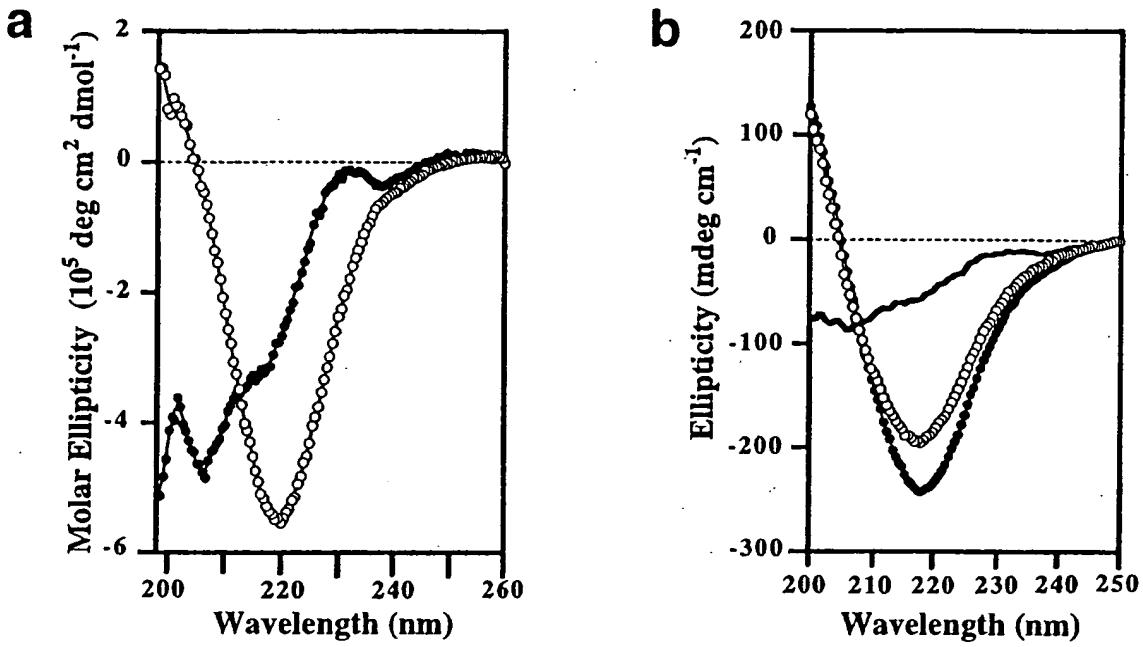


FIG. 2. Far UV CD spectra of Sup35pN in 0.1% TFA/40% acetonitrile. (*a*) Normalized molar CD spectra of a freshly prepared sample (●) and an aged sample (○). (*b*) Ellipticity in a 40 μM aged Sup35pN solution (●) and its sedimented (○; resuspended for the CD measurement in 0.1% TFA/40% acetonitrile) and nonsedimented (solid line) parts after centrifugation at 20,000 × *g*. Small amounts of Sup35pN may have dissociated from the filaments during sedimentation and may contribute to the residual CD signal in the supernatant.

have increased protease K resistance when compared with the control (*ii*) (Fig. 4).

A key experiment in this investigation showed convincingly that seeding with filaments of Sup35pN leads to recruitment of more Sup35pN polypeptide into filaments. After addition of 1% (wt/wt) of the filaments to a freshly prepared solution of

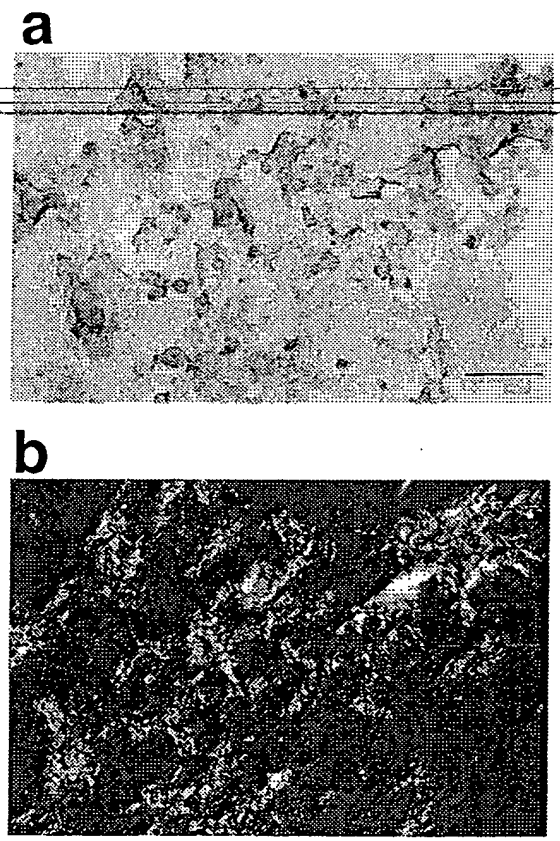


FIG. 3. Photomicrographs of Sup35pN filaments stained with Congo red: (*a*) bright field, (*b*) polarized light. (Bar = 50 μm.)

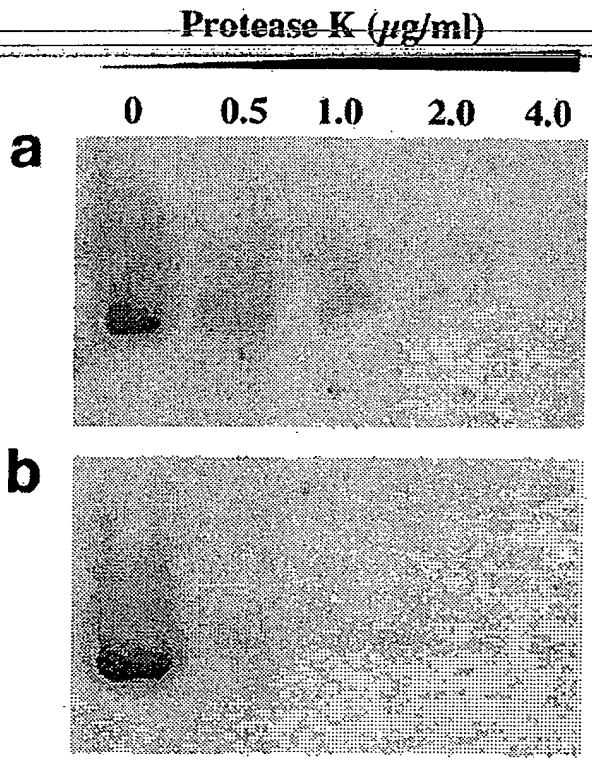


FIG. 4. Protease K resistance of Sup35pN filamentous aggregates (*a*) and of amorphous aggregates (*b*) as determined by SDS/PAGE. The protease K concentration in each reaction is indicated.

BEST AVAILABLE COPY

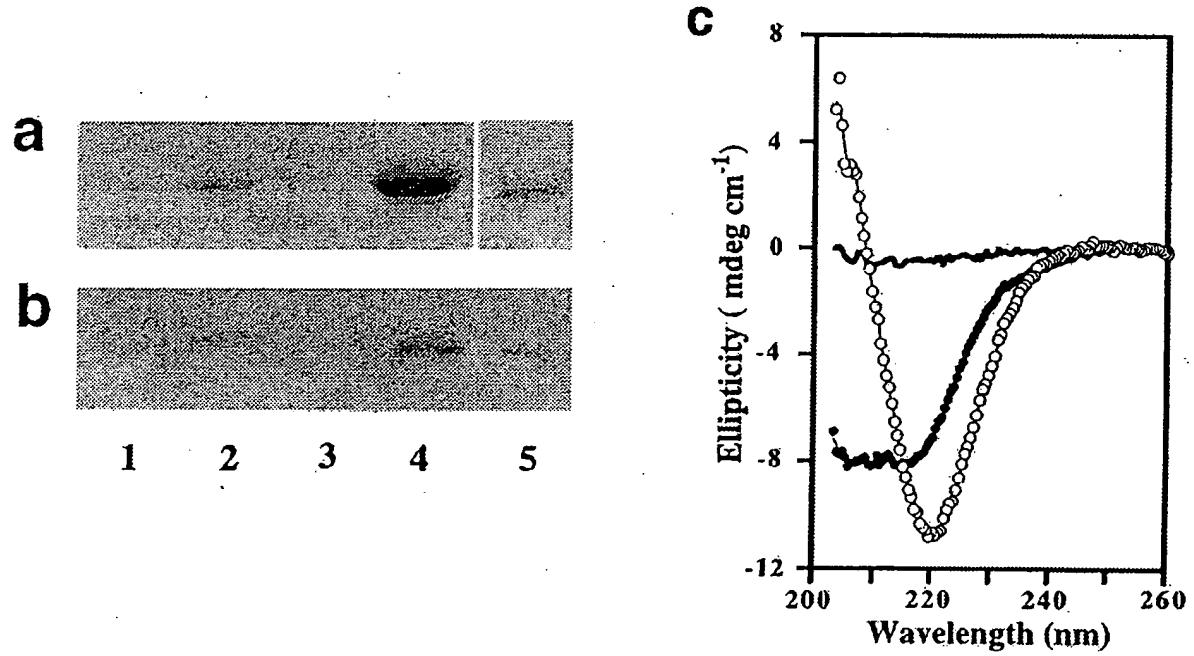


FIG. 5. Gels showing filament growth after seeding in 0.1% TFA/40% acetonitrile (*a*) or in 50 mM sodium phosphate at pH 7.0 (*b*). Lane 1, mock-seeded; lane 2, seeding with amorphous aggregates; lane 3, seeding with GdmCl-denatured filaments; lane 4, seeding with filaments; lane 5, seeds alone directly loaded. (*c*) CD spectra of 5 μM Sup35pN in 50 mM sodium phosphate at pH 7.0: (●) mock-seeded, (○) seeded with filaments; the solid line corresponds to the spectrum of the added seeds.

Sup35pN in 40% acetonitrile/60% H_2O/0.1% TFA at pH 2.0 and overnight incubation at 4°C, large amounts of filamentous species could be isolated by centrifugation (Fig. 5*a*, lane 4) and identified by EM, while a mock-seeded solution yielded no sediments (Fig. 5*a*, lane 1). Formation of new filamentous structures was clearly associated with the addition of filaments, since the seeding activity was lost after denaturation of the filaments by 8 M GdmCl before they were used for seeding (Fig. 5*a*, lane 3). In another control, no increase of the degree of aggregation was seen within 1 day after seeding with the same amount of amorphous Sup35pN aggregates (Fig. 5*a*, lane 2).

Seeding was also observed at pH 7.0 in sodium phosphate buffer without organic cosolvent (Fig. 5*b*, lane 4). The association between seeding activity and presence of the filamentous structure was confirmed using the same controls as described above (Fig. 5*b*, lanes 1–4). This relation was further confirmed by comparison of the CD spectra of seeded and mock-seeded solutions after overnight incubation at 22°C and brief sonication. Fig. 5*c* shows that under these near-physiological solution conditions, seeding did also result in an increased β-type CD signal when compared with the mock-seeded reaction.

DISCUSSION

Previous work by others has established that Sup35 protein extracted from [PSI$^+$] cells forms aggregates that exhibit partial protease resistance and interact with a recombinant Sup35p peptide fragment *in vitro* (12, 13). The presently described ordered filaments, which were prepared from a physiologically relevant peptide (9, 11), retain all the biochemical properties reported for *ex vivo* Sup35p aggregates. Although acidic pH and an organic cosolvent were initially used to suppress amorphous aggregation, once filaments appeared they were stable and capable of further growth also in aqueous solution at neutral pH.

The observed delayed filament formation in the absence of seeding and the seeding behavior of Sup35pN would be compatible with growth kinetics similar to those reported previously for amyloid fibrils, which elongate from critical nucleation sites of which the formation is rate-limiting (19–24). Such kinetics would offer an explanation for major features of [PSI] inheritance: spontaneous [psi$^-$] to [PSI$^+$] reversion is rare because *de novo* formation of an active nucleation site is difficult (10, 25); once such sites exist, propagation of the [PSI] determinant may involve recruiting newly synthesized peptide into filaments, which can release new nucleation sites (12). A similar mechanism was also proposed to rationalize the propagation of prions in mammalian species (refs. 19, 26, and 27; see also refs. 28 and 29 for alternative interpretations), where the infectious form of the prion protein is also an oligomer with amyloid characteristics (30). In this context, it is of further interest that EM studies of aged solutions of the prion-inducing domain of Ure2p (31), which is the cytosolic determinant of [URE3] (3, 7), also revealed filamentous structures (C.-Y.K. and K.W., unpublished work). The fact that mammalian prion proteins Sup35pN and Ure2p, which are all thought to undergo self-propagating conformation changes, are also able to adopt filamentous structures suggests that filament growth by seeding might be a common mechanism for prion propagation.

We thank Dr. A. Jancso for the expression plasmid; Dr. S. te Heesen for the yeast genomic DNA; Dr. G. Frank for peptide sequencing; Drs. S. te Heesen, R. Brunisholz, and Y. Chernoff for technical advice; Drs. R. Glockshuber and C. Weissmann for reading of the manuscript and helpful discussions; and Mrs. E. Ulrich for the careful processing of the manuscript. Financial support was obtained from the Human Frontier Science program, the Schweizerischer Nationalfonds (project 31.32033.91), and the Roche Foundation.

1. Prusiner, S. B. (1996) *Trends Biol. Sci.* **21**, 482–487.
2. Weissmann, C. (1996) *FEBS Lett.* **389**, 3–11.
3. Wickner, R. B. (1994) *Science* **264**, 566–569.
4. Wickner, R. B., Masison, D. C. & Edskes, H. K. (1995) *Yeast* **11**, 1671–1685.

5. Tuite, M. F. (1994) *Nature (London)* **370,** 327–328.
6. Cox, B. S., Tuite, M. F. & McLaughlin, C. S. (1988) *Yeast* **4,** 159–178.
7. Lacroute, F. (1971) *J. Bacteriol.* **106,** 519–522.
8. Stansfield, I., Jones, K. M., Kushnirov, V. V., Dagkesamanskaya, A. R., Poznyakovski, A. I., Paushkin, S. V., Nierras, C. R., Cox, B. S., Ter-Avanesyan, M. D. & Tuite, M. F. (1995) *EMBO J.* **14,** 4365–4373.
9. Ter-Avanesyan, M. D., Dagkesamanskaya, A. R., Kushnirov, V. V. & Smirnov, V. N. (1994) *Genetics* **137,** 671–676.
10. Chernoff, Y. O., Derkach, I. L. & Inge-Vechtomov, S. G. (1993) *Curr. Genet.* **24,** 268–270.
11. Derkatch, I. L., Chernoff, Y. O., Kushnirov. V. V., Inge-Vechtomov, S. G. & Liebman, S. (1996) *Genetics* **144,** 1375–1386.
12. Paushkin, S. V., Kushnirov, V. V., Smirnov, V. N. & Ter-Avanesyan, M. D. (1996) *EMBO J.* **15,** 3127–3134.
13. Patino, M. M., Liu, J. J., Glover, J. R. & Lindquist, S. (1996) *Science* **273,** 622–626.
14. Way, M., Pope, B., Gooch, J., Hawkins, M. & Weeds, A. G. (1990) *EMBO J.* **9,** 4103–4109.
15. Gross, E. (1967) *Methods Enzymol.* **11,** 238–255.
16. Schägger, H. & von Jagow, G. (1987) *Anal. Biochem.* **166,** 368–379.
17. Sipe, J. D. (1992) *Annu. Rev. Biochem.* **61,** 947–975.
18. Glenner, G. G. & Page, D. L. (1976) *Int. Rev. Exp. Pathol.* **15,** 1–92.
19. Jarrett, J. T. & Lansbury, P. T., Jr. (1993) *Cell* **73,** 1055–1058.
20. Lomakin, A., Chung, D. S., Benedek, G. B., Kirschner, D. A. & Teplow, D. B. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 1125–1129.
21. Jarrett, J. T. & Lansbury, P. T., Jr. (1992) *Biochemistry* **31,** 12345–12352.
22. Come, J. E., Fraser, P. E. & Lansbury, P. T., Jr. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 5959–5963.
23. Colon, W. & Kelly, J. W. (1992) *Biochemistry* **31,** 8654–8660.
24. Wood, S. J., Maleeff, B., Hart, T. & Wetzel, R. (1996) *J. Mol. Biol.* **256,** 870–877.
25. Lund, P. M. & Cox, B. S. (1981) *Genet. Res.* **37,** 173–182.
26. Brown, P., Goldfarb, L. G. & Gajdusek, D. C. (1991) *Lancet* **337,** 1019–1022.
27. Caughey, B., Kocisko, D. A., Raymond, G. J. & Lansbury, P. T., Jr. (1995) *Curr. Biol.* **2,** 807–817.
28. Prusiner, S. B. (1991) *Science* **252,** 1515–1522.
29. Wille, H., Zhang, G.-F., Baldwin, M. A., Cohen, F. E. & Prusiner, S. B. (1996) *J. Mol. Biol.* **259,** 608–621.
30. DeArmond, S. J., McKiniey, M. P., Barry, R. A., Braunfeld, M. B., McColloch, J. R. & Prusiner, S. B. (1985) *Cell* **41,** 221–235.
31. Masison, D. C. & Wickner, R. B. (1995) *Science* **270,** 93–95.

72. A. Vortkamp *et al.*, data not shown.
73. Our model suggests that PTHrP represses chondrocyte differentiation, acting on prehypertrophic cells just before the onset of Ihh expression. However, the inhibitory PTHrP signal may also act, to some extent, on the proliferating chondrocytes before the prehypertrophic stage. In the mouse, both antibodies to the receptor and PTH-binding studies suggest that a very low level of PTH/PTHrP receptor may be present on the surface of proliferating cartilage cells even though the levels of PTH/PTHrP receptor mRNA are too low to be detected (*72*).
74. We thank R. L. Johnson for the genomic Ihh probes and sequence information; V. Marigo for the probes for Gli and Ptc; B. R. Olsen for the Col-IX and Col-X probes; B. Houston for the Bmp-6 probes; and T. Woolf, L. Wang, and Ontogeny for providing Sonic hedgehog protein, prepared from an unpublished clone provided by H. Roelink. We are grateful to V. Marigo, J. Capdevilla, E. Laufer, and M. Belliveau for critical discussion of the manuscript and to the members of the Tabin lab for technical advice. A.V. was supported by a fellowship of the Human Frontiers Science Program (HFSP) (LT246/94). Supported by a grant from the HFSP (to C.J.T.) and by National Institutes of Health grants DK47038 (to H.M.K.) and DK4723 to (G.V.S.).

10 April 1995; accepted 20 June 1996

Support for the Prion Hypothesis for Inheritance of a Phenotypic Trait in Yeast

Maria M. Patino, Jia-Jia Liu, John R. Glover, Susan Lindquist*

A cytoplasmically inherited genetic element in yeast, [*PSI*$^+$], was confirmed to be a prionlike aggregate of the cellular protein Sup35 by differential centrifugation analysis and microscopic localization of a Sup35–green fluorescent protein fusion. Aggregation d pended on the intracellular concentration and functional state of the chaperone protein Hsp104 in the same manner as did [*PSI*$^+$] inheritance. The amino-terminal and carboxy-terminal domains of Sup35 contributed to the unusual behavior of [*PSI*$^+$]. [*PSI*$^+$] altered the conformational state of newly synthesized prion proteins, inducing them to aggregate as well, thus fulfilling a major tenet of the prion hypothesis.

Mammalian prions cause devastating neurodegenerative disorders (*1*). Unlike conventional pathogens, they are thought to consist entirely of protein—specifically, a normal nuclear-encoded protein, PrP^C, with an altered "scrapie" conformation, PrP^{Sc} (*1*). The key to prion pathology is thought to be the ability of PrP^{Sc} to induce new PrP^C molecules to adopt the altered structure, producing a protein-conformation cascade that causes the disease and gives rise to new infectious PrP^{Sc}.

A similar explanation can account for the otherwise baffling behavior of two genetic factors in yeast, [*PSI*$^+$] and [*URE3*] (*2*). The [*PSI*$^+$] factor increases translational read-through of all three nonsense codons, and is monitored in the laboratory by omnipotent suppression of nonsense mutations (*3*). Although unlinked to any known nucleic acid, [*PSI*$^-$] behaves as a dominant, cytoplasmically inherited genetic element. It bears an unusual relation to the nuclear-encoded protein Sup35 that is reminiscent of the relation between mammalian prions and nuclear-encoded PrP^C (*1–4*).

Normally, Sup35 is a subunit of the translation-release factor that causes ribosomes to terminate translation at nonsense codons. Release activity maps to the COOH-terminal domain (*5*), which is essential for growth (*6*). Sup35's NH_2-terminal domain is not essential and is required only for the propagation of [*PSI*$^+$] (*6*). Mutations in Sup35 can also cause omnipotent nonsense suppression, but unlike [*PSI*$^+$], the mutant phenotypes exhibit Mendelian inheritance (*3*). Remarkably, transient overexpression of Sup35, or just its NH_2-terminal domain, can induce de novo heritable [*PSI*$^+$] elements (*2*, *6*). Moreover, transient overexpression of the chaperone Hsp104 can restore translational fidelity, heritably converting cells from [*PSI*$^+$] to [*psi*$^-$] (*4*).

These observations argue that [*PSI*$^+$] represents the inheritance of a self-perpetuating alteration in the conformation of Sup35, which is initiated by the NH_2-terminal domain and impairs the ability of the COOH-terminal domain to function in translation. Although this mechanism successfully explains many perplexing genetic observations (*2*, *3*), such a revolutionary model for the inheritance of a phenotypic trait demands the support of direct physical evidence, which we provide here.

Insolubility of Sup35 in [*PSI*$^+$] cells. Isogenic [*psi*$^-$] and [*PSI*$^+$] strains of two different genetic backgrounds (*7*) contained the same quantity of Sup35 and Sup45 (Fig. 1A), the other subunit of the translation-release factor (*8*). Thus, the read-through of nonsense codons in [*PSI*$^-$] cells was not due to reduced accumulation of the termination-factor subunits. Nor was it likely due to posttranslational modification. The migration of the Sup35 and Sup45 proteins from [*PSI*$^+$] cells on high resolution two-dimensional gels was identical to that of the proteins from [*psi*$^-$] cells (*9*).

In contrast, the solubility of Sup35 was very different in [*PSI*$^+$] and [*psi*$^-$] cells. Most Sup35 protein in [*PSI*$^+$] lysates pelleted after centrifugation at 12,000*g*; most remained in the supernatant of [*psi*$^-$] lysates. In [*psi*$^-$] lysates, a substantial fraction of Sup35 remained soluble after centrifugation at 100,000*g*; none remained soluble in [*PSI*$^+$] lysates (Fig. 1B). Similar differences in the solubility of Sup35 in [*PSI*$^+$] and [*psi*$^-$] cells were obtained in the early, mid-, and late log phases of growth as well as in cells in the stationary phase (Fig. 1C) (*10*). No difference in the sedimentation properties of total proteins was detected by Coomassie blue staining, nor did immunostaining show any difference in the sedimentation of Sup45, ribosomal protein L3, or the chaperone proteins Hsp70, Hsp90, and Hsp104 (Fig. 1, B and C) (*10*). High salt (1 M KCl), EDTA (50 mM), and ribonuclease A (400 μg/ml) treatments did not reduce the quantity of Sup35 found in the pellet of [*PSI*$^+$] cells, nor did treatments with nonionic detergent (1% Triton X-100) (*10*). Moreover, like PrP^{Sc} (*1*), the Sup35 protein found in these aggregates was resistant to proteolysis (*11*).

Role of the chaperone Hsp104 in Sup35 aggregation. Overexpression of Hsp104, a protein that promotes the resolubilization and reactivation of heat-damaged proteins (*12*), converts cells from [*PSI*$^+$] to [*psi*$^-$] (*6*). If aggregates of Sup35 reflect the presence of [*PSI*$^+$], Sup35 should return to the soluble state after this conversion. When cells were transformed with a centromeric vector expressing Hsp104 from its own promoter, this was indeed the case (Fig. 1D). [In this and all experiments reported here, the [*PSI*$^+$] and [*psi*$^-$] states were confirmed by plating assays on selective media (Fig. 2) (*13*).] A stronger test of the relation between Sup35 aggregates and [*PSI*$^+$] derives from the ability of transient Hsp104 overexpression to heritably cure

The authors are at the Howard Hughes Medical Institute and the Department of Molecular Genetics and Cell Biology, University of Chicago, 5841 South Maryland Avenue, Chicago, IL 60637, USA.

*To whom correspondence should be addressed.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

cells of [*PSI*$^+$] (6). A galactose-inducible single-copy vector, which provides uniform expression in all cells, converted more than 80% of [*PSI*$^-$] cells to [*psi*$^-$] after 3.5 hours of galactose induction. In converted cells plated to glucose media without continued selection for the Hsp104 plasmid (*13*), Sup35 was found in the soluble fraction, even though Hsp104 was no longer overexpressed (*10*). Similar results were obtained with a glucocorticoid-inducible Hsp104 expression vector (*10*).

Paradoxically, [*PSI*$^+$] cells also convert to [*psi*$^-$] when Hsp104 expression is eliminated by deletion of the *HSP104* gene (*6*). This is surprising because the only previously known function of Hsp104 was to promote the dissolution of aggregates of heat-damaged proteins; these aggregates are maintained in cells with *HSP104* deletions (*12*). When [*PSI*$^-$] cells were converted to [*psi*$^-$] through deletion of *HSP104*, Sup35 was found in the soluble fraction (Fig. 1E). Thus, the behavior of Sup35 aggregates paralleled the behavior of [*PSI*$^-$] and differed from the behavior of heat-damaged aggregates. Together these data strongly support the hypothesis that [*PSI*$^+$]-mediated nonsense suppression is due to a conformational alteration in Sup35 that is self-sustaining as long as Hsp104 is present at its normal concentration.

The sedimentation properties of Sup35 also provide a simple molecular explanation for one of the most perplexing aspects of [*PSI*$^+$] biology—the ability of the element to exist in a cryptic form. For example, when [*PSI*$^+$] cells were transformed with plasmids encoding mutations in the Hsp104 nucleotide-binding domains (NBDs) [either K218T (a Lys to Thr substitution at amino acid 218 in NBD 1) or K620T (an analogous substitution in NBD 2)], the [*PSI*$^+$] phenotype was suppressed but not cured (*4*) (Fig. 2B). That is, the cells did not exhibit nonsense suppression and were unable to grow on selective media, but when the plasmid encoding the mutant Hsp104 protein was lost, [*PSI*$^-$] reappeared and growth on selective media was restored. In contrast, the double mutant K218TK620T cured cells of [*PSI*$^+$]; when the expression plasmid was lost, [*PSI*$^+$]-mediated nonsense suppression was not regained (Fig. 2B).

As shown in Fig. 1F, a greater fraction of Sup35 remained soluble in cells expressing K218T or K620T than in the original [*PSI*$^+$] strain, but most of the the protein remained insoluble (Fig. 1F). Presumably, the increase in soluble Sup35 allowed faithful termination at nonsense codons, but a sufficient quantity of aggregated Sup35 remained to reestablish [*PSI*$^+$] when the plasmid encoding the mutant protein was lost. A smaller fraction of Sup35 was insoluble in the double mutant, and this material was unable to act as a prion-inducing element.

Visualizing prion elements. To monitor [*PSI*$^+$] elements in real time in living cells, we used a green fluorescent protein (GFP) fusion (*14*). The NH_2-terminal prion-deter-

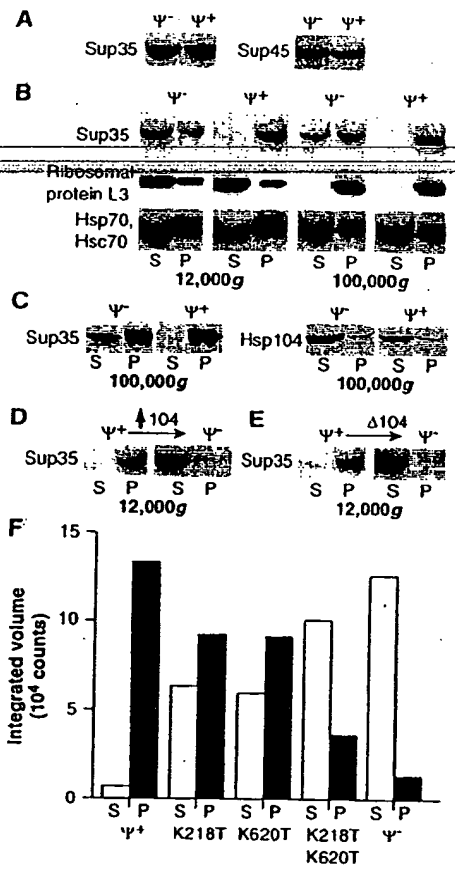


Fig. 1. Sup35 aggregates in [*PSI*$^+$] but not [*psi*$^-$] cells. (**A**) Sup35 and Sup45 accumulation in [*PSI*$^+$] and [*psi*$^-$] cells. Electrophoretically separated total cellular proteins were reacted with antibodies directed against Sup35 and Sup45 (*7, 22, 25*). (**B**) Solubility of Sup35 in mid-log phase cells. Proteins fractionated by centrifugation at 12,000 or 100,000*g* were electrophoretically separated and reacted with antibodies against Sup35 (*22*), ribosomal protein L3 (*25*), and members of the Hsp70 and Hsc70 family [monoclonal antibody (mAb) 7.10]. (**C**) Solubility of Sup35 in stationary-phase cells. Proteins were analyzed as in (B) and reacted with antibodies to Sup35 (*22*) and Hsp104 (*12*). Hsp104 is more readily visualized in stationary-phase cells because Hsp104 concentrations are higher, but similar results were obtained in log-phase cells (*10*). (**D**) Sup35 is soluble in cells overexpressing Hsp104. [*PSI*$^+$] 74D-694 cells (Ψ$^-$) were converted to [*psi*$^-$] (Ψ$^-$) by transformation with a centromeric plasmid pYS104 carrying the wild-type *HSP104* gene, which increases Hsp104 expression two- to threefold as compared with that of wild-type cells (*4*). Fractionated lysates were analyzed as in (B). (**E**) Sup35 is soluble in *HSP104* deletion mutants (Δ104). [*PSI*$^+$] 74D-694 cells (Ψ$^-$) were converted to [*psi*$^-$] (Ψ$^-$) by transformation with a construct [*hsp104::Leu2*] (*4*) that abolishes Hsp104 expression by disrupting the chromosomal *HSP104* gene. Fractionated lysates were analyzed as in (B). (**F**) Solubility of Sup35 in cells harboring cryptic [*PSI*$^+$] elements. [*PSI*$^+$] 74-D694 cells were converted to [*psi*$^-$] by transformation with centromeric plasmids encoding Hsp104 proteins with Lys to Thr substitutions in the first (K218T) or second (K620T) or both (K218TK620T) NBDs (*4*). The parental strain (Ψ$^+$) and its isogenic [*psi*$^-$] derivative (Ψ$^-$) are also shown. Immmunoblots were quantified with ImageQuant software (Molecular Dynamics).

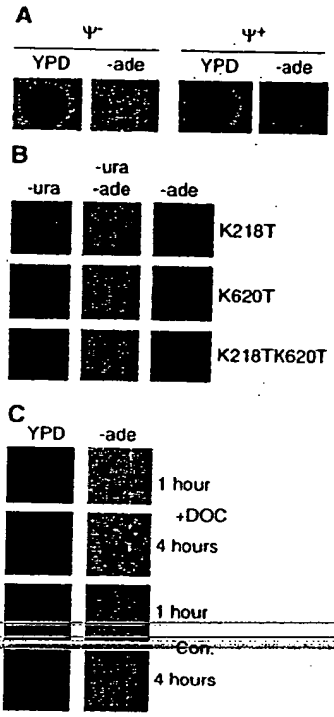


Fig. 2. Analysis of [*PSI*$^-$] propagation by colony formation. (**A**) Read-through of nonsense codons in [*PSI*$^+$] cells detected by the suppression of nonsense mutations. In 74-D694 cells, the suppressible marker is *ade 1-14* (UGA). [*psi*$^-$] cells (Ψ$^-$) do not form colonies on adenine-deficient medium (-ade) and are red on YPD; [*PSI*$^+$] cells (Ψ$^-$) form colonies on adenine-deficient medium and are white on YPD (*4, 13*). (**B**) Cells expressing K218T or K620T Hsp104 mutations carry cryptic [*PSI*$^+$] elements. [*PSI*$^+$] cells analyzed in Fig. 1F were spotted onto plates deficient in uracil (-ura) or adenine (-ade) or both. Growth without uracil forces retention of the plasmid. Growth on adenine-deficient medium requires read-through of the *ade 1-14* UGA codon and the reappearance of [*PSI*$^+$]; it occurs only when cells are allowed to lose the K218T or K602T expression plasmid. (**C**) Transient expression of NPD-GFP induces [*PSI*$^-$]. [*psi*$^-$] 74-D694 cells carrying the *GRE*-regulated NPD-GFP expression plasmid (*15*) were treated with DOC (1 μM) for 1 or 4 hours. Equal numbers of induced (+DOC) and uninduced control (con.) cells were spotted onto YPD and adenine-deficient medium (-ade). [*psi*$^-$] strains underwent conversion to [*PSI*$^+$] after as little as 1 hour of induction, and the extent of conversion increased over time. NPD-GFP did not influence colony formation by [*PSI*$^+$] 74-D694 cells (*16*).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

BEST AVAILABLE COPY

mining domain (NPD) of Sup35 was fused to GFP and placed under the control of the regulatable promoters *CUP1* (inducible with copper) and *GRE* [inducible with 11-deoxycorticosterone (DOC)] (*15*). When NPD-GFP was induced by either copper or DOC, fluorescence was diffusely distributed in two different [*psi*$^-$] strains (*7*). In their isogenic [*PSI*$^+$] derivatives, as soon as fluorescence could be detected it was concentrated in a small number of intense foci (Fig. 3A). When subjected to differential centrifugation, NPD-GFP sedimented with Sup35 in [*PSI*$^+$] lysates but remained in the supernatant of [*psi*$^-$] lysates (*10*). When expressed without the NPD, GFP was diffusely distributed and soluble in both [*psi*$^-$] and [*PSI*$^-$] cells (Fig. 3B). Thus, the coalescence of GFP in [*PSI*$^+$] strains depended on both the attached NPD and the presence of preexisting [*PSI*$^+$] elements.

The Sup35 NPD can induce [*PSI*$^-$] elements in [*psi*$^-$] cells (*6*). In our study, aggregates appeared in a small percentage of the copper-treated [*psi*$^-$] cells after 1 hour of induction (*16*). When plated onto media selective for nonsense suppression but not selective for the NPD-GFP plasmid, heritable [*PSI*$^+$] elements were detected in a similar small percentage of cells (*16*). When the NPD-GFP fusion protein was expressed in [*psi*$^-$] cells at a higher level or for a longer period, bright points of coalescence appeared in a larger fraction of the cells (Fig. 3A) (*16*), and a correspondingly larger fraction showed conversion to [*PSI*$^+$] (Fig. 2C). Intense fluorescent foci were maintained in mother and daughter cells for at least 4 hours after NPD-GFP expression was repressed. In contrast, when NPD-GFP was expressed at high levels in an *HSP104* deletion strain, which cannot propagate [*PSI*$^+$], GFP coalescence was observed in only a few rare cells (*16*). Thus, NPD-GFP coalescence is a marker of the heritable prionlike state of Sup35.

Next we used NPD-GFP to visualize cryptic [*PSI*$^+$] elements in cells expressing mutant Hsp104 proteins. Cells expressing the K218T and K620T proteins exhibited the diffuse fluorescence characteristic of [*psi*$^-$] cells, but many also contained the intense foci characteristic of [*PSI*$^+$] cells (Fig. 3D). In contrast, intense foci were rarely observed in cells expressing the K218TK620T double mutant. Thus, although some Sup35 protein was insoluble in the latter (Fig. 1F), it did not efficiently nucleate the coalescence of newly synthesized NPD-GFP nor the reappearance of [*PSI*$^-$] (Fig. 2B).

Unique properties of Sup35 [*PSI*$^+$] aggregates. To further probe the relation between protein aggregation and prion inheritance, we monitored the behavior of another aggregation-prone GFP protein, a run-on translation product generated by mutation of the termination codon (*15*). This protein (GFP-t) was more variable in expression than NPD-GFP, accumulating in only a fraction of the cells. In these, its distribution varied widely: In groups of budding cells, some individuals exhibited diffuse fluorescence, whereas others showed intense concentrated foci (Fig 3C). Thus, unlike that of NPD-GFP, the distribution pattern of GFP-t was not inherited.

During the formation of aggregates, other amyloids and prions are thought to abandon most of their normal structure. Because GFP fluorescence depends on proper tertiary structure (*17*), some structure must be maintained in the prionlike foci of NPD-GFP. We do not yet know whether the COOH-terminal domain of Sup35 retains its tertiary structure during prion formation, but further experiments demonstrated that it strongly affected the behavior of the NPD.

To determine how Hsp104 affects the solubility of the NH_2- and COOH-terminal domains of Sup35, we expressed them separately in wild-type cells and in *HSP104* deletion mutants (Fig. 4). In wild-type cells, each domain was present in the supernatant and pellet fractions after a 100,000*g* spin. In *hsp104* deletion mutants, the distribution of the COOH-terminal domain was unchanged, but the NH_2-terminal domain was found only in the pellet. Apparently, the NH_2-terminal domain has an intrinsic ability to interact with Hsp104, and through this interaction to undergo a change in state that alters its solubility (*18*). The aggregates formed by the NPD alone, however, behaved like the amorphous aggregates of denatured proteins that accumulate after heat shock and remain insoluble in *HSP104* deletion mutants (*12*). This contrasted with the be-

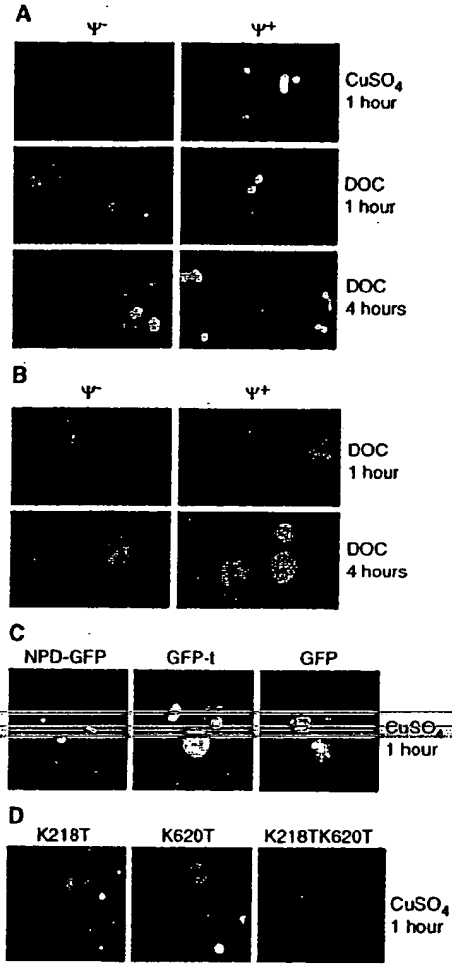


Fig. 3. Visualization of protein aggregates with GFP. (**A**) Diffuse distribution of NPD-GFP in [*psi*$^-$] strains (Ψ$^-$) and coalescence in [*PSI*$^+$] strains (Ψ$^+$). Isogenic [*psi*$^-$] and [*PSI*$^+$] cells (74D-694) were transformed with a *CUP1* or *GRE* NPD-GFP expression plasmid (*15*). $CuSO_4$ (50 μM) or DOC (1 μM) was added to log-phase cultures for 1 or 4 hours. Because our antibodies do not recognize Sup35 protein in its native state, we could not perform colocalization studies. (**B**) GFP is diffusely distributed in all cells. Analysis was as in (A), except that the *GRE* regulated plasmid encoded GFP without the NPD domain. (**C**) The heritability of GFP fluorescence patterns in groups of budding cells. Analysis was as in (A), with plasmids encoding NPD-GFP, GFP-t, and GFP (*15*) under the control of a *CUP1* promoter. (**D**) Cryptic [*PSI*$^-$] elements visualized by NPD-GFP fluorescence. Analysis was as in (A), (*CUP1*-regulated NPD-GFP) in cells transformed with plasmids expressing mutant Hsp104 proteins (see Fig. 1F).

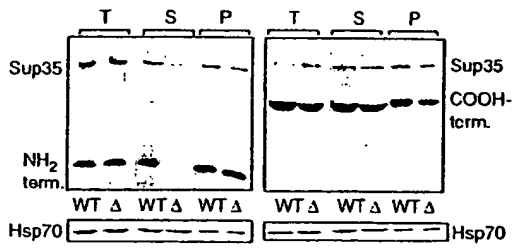


Fig. 4. Hsp104 influences the solubility of the Sup35 NH_2-terminal domain. [*psi*$^-$] cells (74-D694) with an intact (WT) or disrupted (Δ) chromosomal *HSP104* gene transformed with high-copy-number plasmids (*6*) encoding either the Sup35 NH_2-terminal (NH_2-term.) or COOH-terminal (COOH-term.) domain. Lysates were subjected to centrifugation at 100,000*g*. Total lysate (T), supernatant (S), and pellet (P) were analyzed as in Fig. 1, with a polyclonal antiserum to Sup35 and mAb 7.10. Immune complexes were visualized with horseradish peroxidase–conjugated protein A and ECL reagent (Amersham).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

BEST AVAILABLE COPY

havior of the NPD in its normal context, attached to the COOH-terminus, where Hsp104 was actually required for aggregation (Fig. 1D) and, moreover, was required at an intermediate concentration. Thus, the COOH-terminal domain of Sup35 profoundly alters the properties of the NPD and the consequences of its interactions with Hsp104.

The prion hypothesis in yeast. Our data demonstrate that Sup35 undergoes a change in state when cells convert from [*PSI*$^+$] to [*psi*$^-$] and from [*psi*$^-$] to [*PSI*$^+$]. This change involves the disappearance and appearance of a unique heritable aggregate that rapidly captures newly synthesized proteins containing the Sup35 NPD and is governed by the chaperone Hsp104 (Fig. 5). The ability of preexisting [*PSI*$^+$] elements to alter the conformational fate of newly synthesized prion proteins provides direct physical support for the prion hypothesis of [*PSI*$^+$] inheritance (*2*).

Aggregation is a hallmark of the change in state associated with the conversion of mammalian PrP^C to PrP^{Sc} (*19*). The many correlations we observed between the insolubility of Sup35 and the presence of [*PSI*$^+$] demonstrate that aggregation is characteristic of yeast prions as well. However, three findings indicate that [*PSI*$^+$] is more than a simple consequence of protein aggregation. First, a substantial fraction of Sup35 remained insoluble in cells expressing the K218TK620T double mutant, yet this material did not efficiently seed the propagation of [*PSI*$^+$] nor the coalescence of NPD-GFP. Second, when GFP was induced to aggregate through a COOH-terminal extension, the aggregates it formed were not heritable. Third, the aggregates formed by the NPD of Sup35 alone were affected by an *Hsp104* deletion in a different manner than were the [*PSI*$^+$] aggregates of wild-type Sup35. We suggest that the COOH-terminal domain affects the packing of the Sup35 aggregates in a manner that is essential to [*PSI*$^+$] propagation.

Genetic analysis of yeast prionlike elements and the application of GFP fusion protein technology provide a supplement to mammalian investigations that should speed our understanding of self-propagating changes in protein structure and may lead to new approaches for therapeutic intervention in neurodegenerative diseases. But this work, together with work on another such yeast element, [*URE3*] (*2*, *20*), has yet broader implications. The existence of prions—elements of inheritance arising from alternative protein conformations—in both mammals and yeast suggests that they are broadly distributed in nature. In the mammalian brain, a nonmitotic tissue, prions were revealed by their capacity to function as infectious agents; in yeast, they were revealed by their ability to produce heritable changes in phenotype. A wide variety of elusive epigenetic phenomena in other organisms may well prove to depend on the maintenance of alternative protein structures. Finally, because the inheritance of the yeast [*PSI*$^+$] elements depends on Hsp104, a chaperone induced by environmental stress, this phenomenon provides a plausible mechanism for the inheritance of an environmentally acquired characteristic.

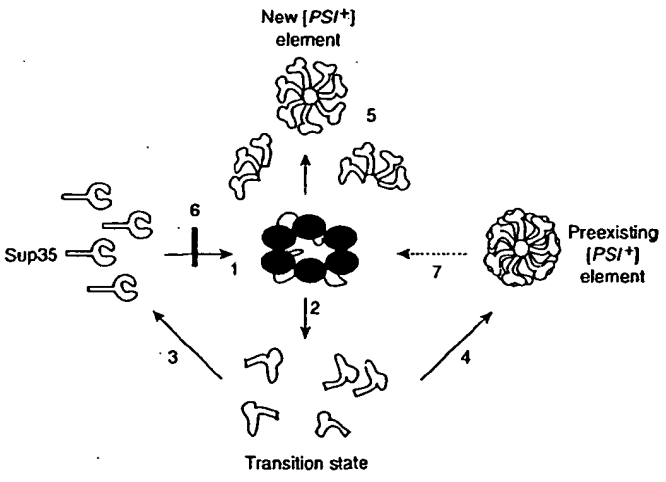


Fig. 5. A model for prion formation in yeast. 1: Newly synthesized Sup35 (white shapes at left) interacts with the chaperone Hsp104 (black ovals at center). 2: Hsp104 helps Sup35 achieve a protein-folding transition state that is required for prion formation but is inherently unstable. 3: In the absence of [*PSI*$^+$], Sup35 reverts to its normal functional state. 4: Preexisting [*PSI*$^+$] elements capture and stabilize transition-state conformers; Sup35 is sequestered from translation and unfaithful termination leads to nonsense suppression. 5: Transient overexpression of Sup35 nucleates prions de novo because the high concentration of transition-state conformers increases the likelihood of stabilizing intermolecular interactions (*23*). 6: In the absence of Hsp104, the transition state is difficult to attain and prions cannot be perpetuated. 7: Overexpression of Hsp104 might disturb the equilibrium in several ways: Hsp104 might bind prion-state conformers and disaggregate them; rebind monomers, reducing their ability to be captured by [*PSI*$^+$] elements; or reduce the local concentration of transition-state conformers because they are dispersed in association with larger numbers of Hsp104 (*24*).

REFERENCES AND NOTES

1. S. B. Prusiner, *Annu. Rev. Microbiol.* **48**, 655 (1994); C. Weissman, *Trends Cell Biol.* **4**, 10 (1994).
2. R. B. Wickner, *Science* **264**, 566 (1994); ———, D. C. Masison, H. K. Edskes, *Yeast* **11**, 1671 (1995).
3. B. S. Cox, M. F. Tuite, C. S. McLaughlin, *Yeast* **4**, 159 (1988); S. Lindquist *et al.*, *Cold Spring Harbor Symp. Quant. Biol.* **60**, 451 (1995); M. Tuite and S. Lindquist, *Trends Genet.*, in press.
4. Y. O. Chernoff, S. L. Lindquist, B.-i. Ono, S. G. Inge-Vechtomov, S. W. Liebman, *Science* **268**, 880 (1995).
5. I. Stansfield *et al.*, *EMBO J.* **14**, 4365 (1995).
6. M. D. Ter-Avanesyan *et al.*, *Mol. Microbiol.* **7**, 683 (1993); Y. O. Chernoff, I. L. Derkach, S. G. Inge-Vechtomov, *Curr. Genet.* **24**, 268 (1993). S. M. Doel *et al. Genetics* **137**, 659 (1994); M. D. Ter-Avanesyan *et al.*, *ibid.*, p. 671.
7. Shown in Fig. 1A is *Saccharomyces cerevisiae* strain 74D-694 (*MATa, ade1 his3 leu2 trp1 ura3*) [*PSI*$^+$] and [*psi*$^-$] derivatives. Not shown is strain D$^-$ 142-1A (*4*).
8. G. Zhouravleva *et al.*, *EMBO J.* **14**, 101 (1995).
9. The first dimension of the gel was performed with the Immobiline Dry Strip system (linear gradient, pH 4 to 7) and the second was performed on an ExcelGel SDS (gradient 8 to 18%) with a Multiphor II apparatus. For cell lysis (of strain 74D-694; [*PSI*$^+$] and [*psi*$^-$]), cells were grown on yeast extract, peptone, and dextrose (YPD) to 5×10^6 cells per milliliter and electrophoresis was done according to the manufacturer's recommendations (Pharmacia Biotech). Total proteins were stained with Coomassie blue, transferred to Immobilon filters, and reacted with antibodies to Sup35 (peptide amino acids 37 to 151). Immune complexes were visualized as described in Fig. 4 (M. M. Patino and S. Lindquist, data not shown).
10. M. Patino and S. Lindquist, data not shown.
11. When cell lysates were treated with proteinase K, the Sup35 protein of [*PSI*$^+$] cells was more resistant to digestion than was that of [*psi*$^-$] cells, but unlike the case with PrP^{Sc}, no specific protease-resistant fragments of Sup35 were detected (M. Patino, S. Lindquist, Y. Chernoff, unpublished data).
12. D. A. Parsell, A. S. Kowal, M. A. Singer, S. Lindquist, *Nature* **372**, 475 (1994).
13. In strain 74D-694, the [*PSI*$^+$] suppressible marker is *ade 1-14*, a UGA codon. [*PSI*$^+$] cells produce white colonies on complete media (YPD) and grow without adenine supplementation; [*psi*$^-$] cells produce red colonies and require adenine. Cryptic [*PSI*$^+$] elements in 74D-694 cells expressing Hsp104 K218T or K620T proteins from *URA* vectors were detected by simultaneous platings on media (i) with adenine and without uracil (all cells grew), (ii) with neither uracil or adenine (maintenance of the plasmid abrogated [*PSI*$^+$]-mediated nonsense suppression and cells did not grow), and (iii) with uracil and without adenine (the small percentage of cells that had spontaneously lost the plasmid at the time of plating recovered [*PSI*$^+$]-mediated nonsense suppression and grew). Cells expressing the double mutant could not recover [*PSI*$^-$] and did not grow on medium with uracil and without adenine.
14. A. D. Cubitt *et al.*, *Trends Biochem. Sci.* **20**, 448 (1995); S. R. Kain *et al.*, *Biotechniques* **10**, 650 (1995).
15. To create GFP-t, a DNA segment encoding GFP but lacking the natural termination codon was amplified with primers 5′-CGGATCCCATATGAGTAAAGGAGAAGAACT$^-$-3′ and 5′-CTCGAGCTCACTAGTTTTGTATAGTTCATCATCCATGCC-3′. The polymerase chain reaction (PCR) product was subcloned into plasmid p2UGpd (pRS426) with GPD promoter (D. Nathan and S. Lindquist, unpublished data) as a Bam HI–Sac I fragment generating the plasmid p2UGFP-t. The complete open reading frame extends into the vector sequence and encodes an aggregation-prone GFP with a 46-residue COOH-terminal extension with the sequence TSELQLLFPLVRVNCALGVIMVIAVSCVKLLSAHNSTQHTSRKHKV-COOH (*21*). To create an unfused GFP, a termination codon was regenerated by cutting p2UGFP-t with Spe I, filling in the overhanging ends, and re-ligating (creating the plasmid p2UGFP). To create NPD-GFP, the DNA encoding the NPD of Sup35 was amplified from pEMBL-SUP2 (*6*) with primers 5′-CGCGGATCCATGTCGGATTCAAACC-3′ and 5′-CG-

BEST AVAILABLE COPY

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

CGGATCCATCGTTAACAACTTCG-3′ and subcloned as a Bam HI fragment into p2UGFP. The resulting construc encodes the NH_2-terminal 253 residues of Sup35 fusec to GFP. Each of these constructs (GFP-t, GFP, and NPD-GFP) was amplified by PCR and subcloned into p2UG [M. Schena, D. Picard, K. R. Yamamoto, *Methods Enzymol.* **194**, 389 (1991)], conferring DOC-inducible expression, and also into pCLUC [D. J. Thiele, *Mol. Cell. Biol.* **8**, 745 (1988)] for copper-inducible expression. Fidelity of constructs was confirmed by dideoxy nucleoside triphosphate sequencing, and the mobility of the expressed proteins was determined by protein immunoblot analysis.

16. J.-J. Liu and S. Lindquist, data not shown.
17. W. W. Ward, C. W.. Cody, R. C. Hart, M. J. Cormier, *Photochem. Photobiol.* **31**, 611 (1980); W. W. Ward and S. H. Bokman, *Biochemistry* **21**, 4535 (1982);.
18. The ability of the NH_2-terminal domain to stimulate Hsp104 adenosine triphosphatase activity suggests that this interaction is direct (E. Schirmer and S. Lindquist, unpublished data).
19. P. T. Lansbury Jr. and B. Caughey, *Curr. Biol.* **2**, 1 (1995); B. Coughey *et al.*, *ibid.*, p. 807.
20. D. C. Masison and R. B. Wickner, *Science* **270**, 93 (1995).
21. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
22. Isogenic [*PSI*$^-$] and [*psi*$^-$] cells were grown to a density of ~5×10^6 cells per milliliter in YPD. Cells were suspended in 50 mM tris-HCl (pH 7.5), 5 mM $MgCl_2$, 10 mM KCl, 0.1 mM EDTA, 1 mM dithiothreitol, cycloheximide (100 μg/ml), 1 mM benzamidine, 2 mM phenylmethylsulfonyl, leupeptin (10 μg/ml), pepstatin A (2 μg/ml), and ribonuclease A (100 μg/ml) and disrupted with glass beads at 4°C. Proteins resolved by SDS–polyacrylamide gel electrophoresis were transferred to Immobilon membranes (Millipore) and reacted with an antiserum against amino acids 137 to 151 of Sup35 or against Sup45 (gift of M. Tuite), followed by ^{125}I-conjugated protein A (ICN Pharmaceuticals), and then exposed to a PhosphorImager screen (Molecular Dynamics).
23. These interactions may be facilitated by the simultaneous binding of several Sup35 proteins to an Hsp104 hexamer or by rapid sequential binding and release of individual conformers in its immediate vicinity. [*PSI*$^+$] is drawn as an ordered aggregate, with reference to a model for mammalian prion formation (*19*). [*PSI*$^+$] aggregates have special properties, but we do not yet know if they form an ordered structure.
24. In preliminary experiments, GFP-marked prions do not disaggregate rapidly when Hsp104 is overexpressed (J.-J. Liu, unpublished data). Given their size, this is not surprising. However, because the prion assay relies on colony formation, and because Hsp104 is long-lived, it is possible that overexpression of Hsp104 simply prevents new prion conformers from joining the prion while preexisting prions are diluted by cell division.
25. We thank Y. Chernoff, S. Liebman, I. Kerkatch, and M. F. Tuite for helpful discussions; M. F. Tuite for antibodies to Sup45; J. Warner for antibodies to ribosomal protein L3; M. D. Ter-Avananesyan and V. V. Kushnirov for plasmids and polyspecific Sup35 antisera; N. Patel for help with figures; and M. Singer, S. Rutherford, and S. K. DebBurman for comments on the manuscript. Supported by a grant from the National Institute of General Medical Sciences (NIH grant GM25874), and the Howard Hughes Medical Institute. J.-J.L. was supported by The Markey Program.

10 June 1996; accepted 5 July 1996

Discover a new sequence.

Visit the SCIENCE On-line Web site and you just may find the key piece of information you need for your research. The fully searchable database of research abstracts and news summaries allows you to look through current and back issues of SCIENCE on the World Wide Web. Tap into the sequence below and see SCIENCE On-line for yourself.

NEW URL

http://www.sciencemag.org

SCIENCE

BEST AVAILABLE COPY

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

amyloid solubility / plaque formation

BIOCHIMICA et BIOPHYSICA ACTA

GENERAL SUBJECTS

LIBRARY
RECEIVED
NOV 3 1965
U.S. DEPARTMENT OF AGRICULTURE
BELTSVILLE BRANCH

INTERNATIONAL WEEKLY OF BIOCHEMISTRY AND BIOPHYSICS
REVUE INTERNATIONALE HEBDOMADAIRE DE BIOCHIMIE ET DE BIOPHYSIQUE
INTERNATIONALE WOCHENSCHRIFT FÜR BIOCHEMIE UND BIOPHYSIK

EDITORIAL BOARD

E. Antonini (Rome)
S. Bergström (Stockholm)
J. D. Bernal (London)
J. Brachet (Brussels)
R. O. Brady (Bethesda)
A. E. Braunstein (Moscow)
D. H. Brown (St. Louis)
T. Bücher (Munich)
H. Chantrenne (Brussels)
E. Chargaff (New York)
J. A. Cohen (Rijswijk)
C. F. Cori (St. Louis)
P. Desnuelle (Marseilles)
E. Dicsfalusy (Stockholm)
A. Engström (Stockholm)
E. F. Gale (Cambridge)
M. Gruber (Groningen)
M. Grunberg-Manago (Paris)
D. J. Hanahan (Seattle)
O. Hayaishi (Kyoto)
P. J. Heald (London)
R. W. Holley (Ithaca)
V. Jagannathan (Poona)
A. Katchalsky (Rehovoth)
R. D. Keynes (Babraham)
M. Klingenberg (Marburg/Lahn)
H. L. Kornberg (Leicester)
D. E. Koshland (Upton)
M. Kotani (Tokyo)
H. A. Krebs (Oxford)
L. F. Leloir (Buenos Aires)
F. Lynen (Munich)
B. G. Malmström (Göteborg)
G. F. Marrian (London)
A. Meister (Boston)
J. Monod (Paris)
A. Monroy (Palermo)
A. Neuberger (London)
H. Neurath (Seattle)
A. J. Oparin (Moscow)
L. E. Orgel (San Diego)
M. Ottesen (Copenhagen)
J. Th. G. Overbeek (Utrecht)
A. B. Pardee (Princeton)
G. N. Ramachandran (Madras)
E. R. Redfearn (Leicester)
P. Reichard (Stockholm)
J. Roche (Paris)
D. Rosen (London)
C. Sadron (Strasbourg)
G. Schramm (Tübingen)
M. Sela (Rehovoth)
E. C. Slater (Amsterdam)
J. D. Smith (Cambridge)
P. Strittmatter (St. Louis)
H. Tamiya (Tokyo)
E. L. Tatum (New York)
J. B. Thomas (Utrecht)
A. W. K. Tiselius (Uppsala)
P. R. Vagelos (Bethesda)
L. L. M. Van Deenen (Utrecht)
A. Von Muralt (Bern)
E. C. Webb (Brisbane)
H. H. Weber (Heidelberg)
C. Weissman (New York)
R. Wurmser (Paris)
R. W. G. Wyckoff (Tucson)

ELSEVIER PUBLISHING COMPANY

AMSTERDAM

Biochim. Biophys. Acta, Vol. 104, No. 2, p. 317—636, July 8, 1965

Completing Vol. 104 (1965)

27th Issue f 1965

BEST AVAILABLE COPY

Vol. 194, No. 3, 1993 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
August 16, 1993 Pages 1380-1386

MOLECULAR CHARACTERISTICS OF A PROTEASE-RESISTANT, AMYLOIDOGENIC AND NEUROTOXIC PEPTIDE HOMOLOGOUS TO RESIDUES 106-126 OF THE PRION PROTEIN

Carlo Selvaggini[1], Luca De Gioia[2], Laura Cantù[3], Elena Ghibaudi[1], Luisa Diomede[1], Flavio Passerini[1], Gianluigi Forloni[1], Orso Bugiani[4], Fabrizio Tagliavini[4] and Mario Salmona[1]*

[1]Istituto di Ricerche Farmacologiche "Mario Negri", Via Eritrea 62, 20157 Milano, Italy

[2]Dipartimento di Chimica Inorganica, Metallorganica ed Analitica, Centro CNR, Università di Milano, Via Venezian 21, 20133 Milano, Italy

[3]Study Center for the Functional Biochemistry of Brain Lipids, Department of Chemistry and Biochemistry, Medical School, University of Milano, Milano, Italy

[4]Istituto Neurologico "Carlo Besta", Via Celoria 11, 20133 Milano, Italy

Received June 30, 1993

In the prion-related encephalopathies the prion protein is converted to an altered form, known as PrP^{Sc}, that is partially resistant to protease digestion. This abnormal isoform accumulates in the brain and its protease-resistant core aggregates extracellularly into amyloid fibrils. We have investigated the conformational properties, aggregation behaviour and sensitivity to protease digestion of a synthetic peptide homologous to residues 106-126 of human PrP, which was previously found to form amyloid-like fibrils *in vitro* and displayed neurotoxic activity toward primary cultures of rat hippocampal neurons. A scrambled sequence of peptide PrP 106-126 was used as a control. By circular dichroism, PrP 106-126 exhibited a secondary structure composed largely of β-sheet, whereas the scrambled sequence of PrP 106-126 showed a random coil structure. The β-sheet content of PrP 106-126 was much higher in 200 mM phosphate buffer at pH 5.0 than in the same buffer at pH 7.0. Laser light scattering analysis showed that PrP 106-126 aggregated immediately after dissolution in 20 mM or 200 mM phosphate buffer, pH 5.0 and 7.0, whereas scrambled PrP 106-126 did not. PrP 106-126 aggregates had an average hydrodinamic diameter of 100 nm and an average molecular weight of $12 \times 10^6 \pm 30\%$ Daltons, corresponding to the aggregation of 6000 ± 30% molecules. Peptide PrP 106-126 showed partial resistance to digestion with Proteinase K and Pronase, whereas scrambled PrP 106-126 was completely degraded by incubation with the enzymes at 37 °C for 30 minutes. © 1993 Academic Press, Inc.

The prion diseases are a group of genetic and/or transmissible neurodegenerative disorders that includes scrapie of sheep and goat, spongiform encephalopathy of cattle, and Creutzfeldt-Jakob

* To whom reprint requests should be addressed at: Istituto di Ricerche Farmacologiche "Mario Negri", Laboratory for Enzyme Research, Via Eritrea 62, 20157 Milano, Italy.

0006-291X/93 $4.00
Copyright © 1993 by Academic Press, Inc.
All rights of reproduction in any form reserved.

BEST AVAILABLE COPY

or Gerstmann-Sträussler-Scheinker (GSS) diseases of humans (1). In spite of clinical and pathological differences, all these conditions have a common feature, i.e. the conversion of a 33-35 kDa sialoglycoprotein designated cellular prion protein (PrP^c) into an abnormal isoform known as scrapie prion protein (PrP^{Sc}) (1). The transformation of PrP^c to PrP^{Sc} appears to be a post-translational event (2, 3, 4), but the precise nature of the protein modification is not known. A distinctive feature of PrP^{Sc} is partial resistance to Proteinase K digestion in conditions where PrP^c is completely degraded (5, 6, 7).

In scrapie and related diseases, the protease-resistant core of PrP^{Sc} corresponds to residues 90-231; it has a molecular weight of 27-30 kDa and is intrinsically able to polymerize into rod-shaped particles with the tinctorial, optical and ultrastructural properties of amyloid fibrils (4, 5, 6, 8, 9). Infrared spectroscopy shows that this PrP fragment (designated PrP 27-30) has an extensive β-sheet secondary structure, as expected for amyloid proteins (10). Thus, the conversion of PrP^c to PrP^{Sc} implies a protein modification that increases the molecule's potential for amyloid fibril formation.

The prototype of PrP amyloidosis is GSS disease (11, 12). The amyloid protein, characterized so far only in patients from the Indiana kindred of GSS (13, 14), is an 11 kDa PrP fragment spanning residues 50 to ca. 150 (15). Using synthetic peptides homologous to consecutive segments of this fragment, we recently found that the most highly amyloidogenic region corresponds to residues 106-126 (16, 17). PrP 106-126 fibrils had a diameter of 4-8 nm and ranged in length from 0.1 to 2 μm; they showed green birefringence under polarized light after Congo red staining and X-ray diffraction patterns similar to those of native amyloid fibrils (16, 17). Chronic application of this peptide to primary rat hippocampal cultures caused neuronal death by apoptosis, suggesting that cerebral accumulation of PrP^{Sc} and PrP amyloid might play an important role in the neuronal degeneration seen in prion-related encephalopathies (18).

In the present study we examined the conformational properties and aggregation behaviour of peptide PrP 106-126 by circular dichroism and laser light scattering. We also investigated the sensitivity of PrP 106-126 aggregates to protease digestion. The results were compared with those obtained with a peptide corresponding to a scrambled sequence of PrP 106-126, which was unable to form fibrils in aqueous solutions and had no neurotoxic activity *in vitro* (18).

Materials and Methods

Peptide Synthesis

The peptides PrP 106-126 (KTNMKHMAGAAAAGAVVGGLG) and scrambled PrP 106-126 (NGAKALMGGHGATKVMVGAAA) were synthesized using solid phase chemistry by a 430A Applied Biosystems instrument. FMOC (9-fluorenylmethoxy carbonyl) was used as the protective group for aminic residues, and 1-hydroxybenzotriazole, 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate and N,N-dicyclohexyl-carbodiimide were used as activators of carboxylic residues. Peptides were cleaved from the resin with phenol/thioanisole/ethanedithiol/trifluoroacetic acid (TFA) and precipitated with cold diethylether. Crude peptides were washed several times with diethylether and purified by reverse-phase HPLC (Model 243, Beckman Instruments Inc., Palo Alto, California, USA) on a Delta-Pak C18 column (19 x 300 mm, 300 Å pore size, 15 μm particle size, Nihon Waters, Tokyo, Japan). The purity and composition of peptides were determined by analytical reverse-phase HPLC, capillary electrophoresis (Quanta 4000 -Millipore, Bedford, MA, USA), and amino acid sequencing (6600 Prosequencer - Milligen, Bedford, MA, USA). Purity was greater than 95% for both peptides.

BEST AVAILABLE COPY

Circular Dichroism

For circular dichroism spectroscopy 5-10 mg of either peptide were dissolved in 3-5 ml of 200 mM phosphate buffer, pH 5.0 or 7.0. The spectra were recorded in quartz cells with an optical path of 0.01 or 0.1 cm using a Jasco J-500 dichrograph. A scan speed of 10 nm/min was used, averaging the data over five measurements.

Mean residue ellipticities were calculated using the following equation

$$(\theta)_M = A \times 3300 \times M / C \times l$$

where A = observed dichroic absorbance, l = path length in cm, C= concentration of the peptide in g/L, and M= mean residue weight.

The spectra were qualitatively analyzed to assess the secondary structure patterns according to Curtis Johnson (19).

Laser Light Scattering

The equipment for laser light scattering measurements was described in detail elsewhere (20, 21). Briefly, it includes an argon ion laser, λ=514.5 nm, a temperature controlled cell and a digital correlator. The absolute intensity calibration was done according to Degiorgio et al. (21). Peptides PrP 106-126 and scrambled PrP 106-126 were dissolved in 20 or 200 mM phosphate buffer, pH 5.0 or pH 7.0, and in MilliQ water at a concentration of 10 - 150 µM. The buffers were filtered before use through a 0.2 µm polycarbonate filter and the final solutions were let into the scattering cell through a 5 µm filter. At the end of the measurements, peptide solutions were analyzed by reverse-phase HPLC to establish their exact concentration.

Protease Digestion

Peptides PrP 106-126 and scrambled PrP 106-126 were dissolved in the reaction buffer (10 mM Tris-HCl, 1 mM $CaCl_2$, 0.5% SDS pH=7.8) at a concentration of 1 mg/ml and incubated for 30 minutes at 37°C with Proteinase K or Pronase, at enzyme-to-substrate ratios of respectively 1:20 and 1:1 (w/w). Proteolysis was terminated adding EGTA (5 mM, final concentration) to samples treated with Proteinase K or rapidly freezing (liquid nitrogen) samples treated with Pronase. Control samples were run in the same conditions except for the absence of proteolytic enzymes. After incubation all the samples were kept at -80 °C until analysis.The degree of enzymatic digestion of peptides PrP 106-126 and scrambled PrP 106-126 was determined by reverse phase HPLC using MilliQ water + 0.1% TFA (A) and acetonitrile + 0.08% TFA (B) as eluents, following a linear gradient from 100% A to 30% A in 23 minutes. The retention time for both PrP 106-126 and scrambled PrP 106-126 was 18.5 minutes.

Results

Peptide PrP 106-126 has a high content of β-sheet in aqueous solutions. Figure 1 reports the CD spectra of the peptides PrP 106-126 and scrambled PrP 106-126 recorded in 200 mM phosphate buffer, at pH 7.0 and pH 5.0. The spectrum of PrP 106-126 at pH 7.0 (Panel A) showed a weak negative band around 218 nm and a broad negative minimum at frequencies below 200 nm, while the spectrum of scrambled PrP 106-126 exhibited a prominent negative absorption at 198 nm.

Qualitative analysis of the CD spectroscopic features suggests that scrambled PrP 106-126 does not exist in a preferred conformation but can be considered in a random coil state because of a very pronounced negative CD band at 198 nm. By contrast, the CD features of the peptide PrP 106-126, whose characteristic random coil band at 198 nm was not very strong and whose negative absorption at 216 nm became significant, suggests a mixture of random coil and ß-sheet conformation. Figure 1 (Panel B) documents the CD spectra of both peptides at pH 5.0. A higher content of ß-sheet-like structures and possibly some degree of ß-turn structure was observed for PrP 106-126. Its spectrum was characterized by a negative absorption at 218 nm and a positive band at wavelengths below 207 nm. In contrast the spectrum of scrambled PrP 106-126 showed again the characteristic strong negative CD absorption at 198 nm which is indicative of a completely random coil structure.

BEST AVAILABLE COPY

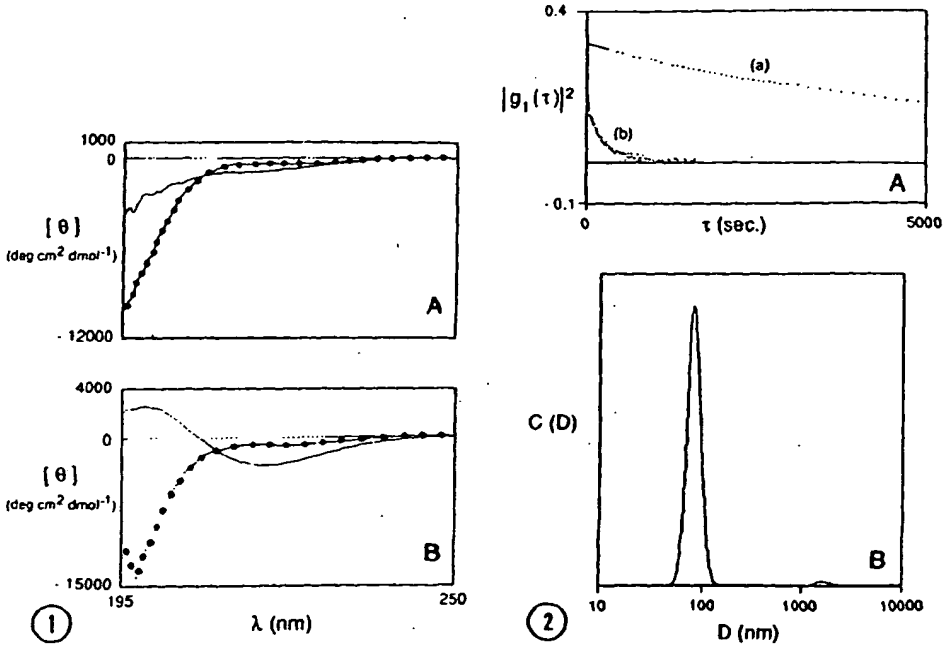


Figure 1. CD spectra of peptides PrP 106-126 (———) and scrambled PrP 106-126 (--•--•--•--) in 200 mM sodium phosphate at pH 7.0 (Panel A) and pH 5.0 (Panel B).

Figure 2. **Panel A.** Intensity correlation functions of peptides PrP 106-126 (35 μM, curve a) and scrambled PrP 106-126 (35 μM, curve b) in 200 mM sodium phosphate, pH 5.0. **Panel B.** Size distribution of particles present in a 35 μM solution of peptide PrP 106-126 in 200 mM sodium phosphate, pH 5.0.

Peptide PrP 106-126 forms high molecular weight aggregates. Figure 2 (Panel A) reports the scattered intensity correlation function of a 35 μM solution of PrP 106-126 (curve a) and scrambled PrP 106-126 (curve b) dissolved in 200 mM phosphate, pH 5.0. Nonlinear analysis of the scattered intensity correlation function indicated that PrP 106-126 (curve a) forms aggregates with wide size distribution, the average hydrodynamic diameter being 100 nm; in addition, the peptide solution contained a small amount of very large particles, whose average hydrodynamic diameter was between 1000 and 2000 nm (Figure 2, Panel B). The scrambled peptide solution contained only small particles, as indicated by the rapid decay of curve b, meaning that aggregates do not exceed a few molecules.

The contribution to the scattered intensity at different observation angles due to the main distribution of aggregates was evaluated and the extrapolated value at $\theta=0°$ was used to determine their average molecular weight. This was $12 \times 10^6 \pm 30\%$ Dalton, corresponding to an aggregation number of 6000 ± 30%. Measurements were also made on PrP 106-126 in deionized water. Aggregates were much smaller, with an hydrodynamic diameter of about 20 nm and an average molecular weight of $0.33\ 10^6 \pm 30\%$ Dalton, corresponding to an aggregation number of about 165± 30%.

No interactive behaviour among aggregates was observed at any concentration and ionic strength ($K_I=K_D=0$) (data not shown). The scattering features of PrP 106-126 were not time-

BEST AVAILABLE COPY

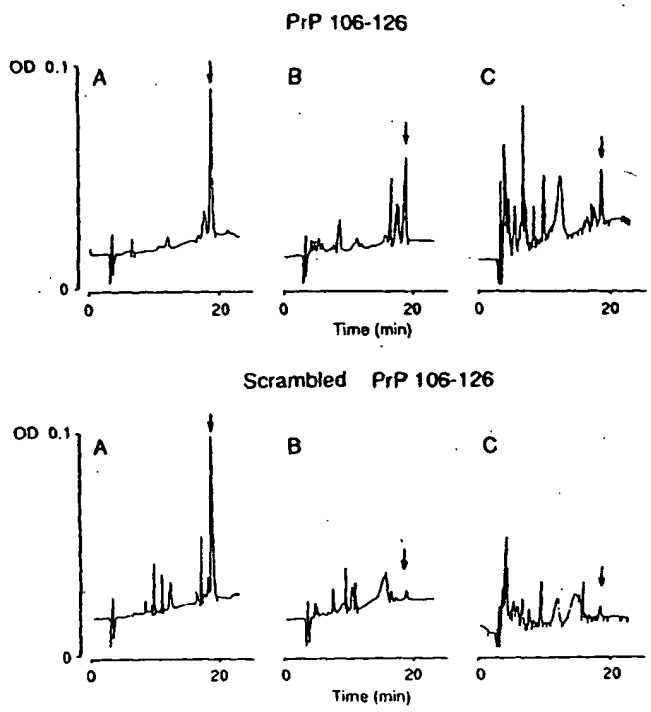


Figure 3. HPLC profile of a solution of 1 mg/ml of peptides PrP 106-126 and scrambled PrP 106-126 after incubation at 37 °C for 30 minutes with Proteinase K (50 µg/ml) or Pronase (1 mg/ml).
The arrow indicates the peak corresponding to the retention times of PrP 106-126 or scrambled PrP 106-126. A. Control. B. Proteinase K. C. Pronase.

dependent, indicating that aggregation occurs immediately after the dissolution of the peptide, within the minute of dead time between sample preparation and laser light scattering measurement. A similar aggregation pattern was observed for the peptide PrP 106-126 when the experiments were performed at pH 7.0 (data not shown).

PrP 106-126 aggregates are partially resistant to Proteinase K and Pronase digestion. Figure 3 reports the HPLC chromatograms showing the relative level of degradation of the peptides PrP 106-126 and scrambled PrP 106-126 after 30 minutes of incubation at 37 °C in the absence of proteases (Panels A) and in the presence of Proteinase K (Panels B) or Pronase (Panels C). In both cases only 60% of PrP 106-126 was hydrolyzed, whereas scrambled PrP 106-126 was degraded more than 95% from its initial level. Time-course experiments showed that PrP 106-126 was not degraded further even after 240 minutes treatment with either enzyme, while scrambled PrP 106-126 was almost completely degraded after only 30 seconds of incubation with Pronase (data not shown).

Discussion

Previous studies with synthetic peptides homologous to consecutive segments of the amyloid protein purified from GSS brains have found that a peptide extending from residue 106 to

BEST AVAILABLE COPY

residue 126 of human PrP is highly capable of assembling into fibrils which display ultrastructural features, tinctorial properties and X-ray diffraction patterns similar to those of *in situ* amyloid (16, 17). We found that the chronic application of this peptide to rat hippocampal cultures caused neuronal death by apoptosis (18). Based on these findings, which extended the observation reported by Gasset et al. (23) that a synthetic peptide homologous to residues 109-122 of Syrian hamster PrP forms amyloid-like fibrils *in vitro* and exhibits a secondary structure composed largely of β-sheet, we postulated that the PrP sequence 106-126 might play a key role in amyloid formation and in the nerve cell degeneration occuring in prion-related encephalopathies.

The aim of the present investigation was to further define the physico-chemical properties of this peptide and to elucidate its behaviour in aqueous media.

The molecular structure of PrP 106-126 consists of a polar head at its N terminal part (KTNMKHM-) and a long hydrophobic core (-AGAAAAGAVVGGLG). This clear cut division is absent in the scrambled peptide, which contains alternate hydrophilic and hydrophobic residues, minimizing the abrupt change in the chemico-physical properties between the hydrophilic and hydrophobic regions of PrP 106-126.

The CD spectrum of PrP 106-126 at pH 7.0 indicated the presence of a combination of ß-sheet and random coil structures, while scrambled PrP 106-126 showed a random coil conformation. Conversely, a more organized secondary structure was evident from experiments with PrP 106-126 at pH 5.0, with the presence of extensive ß-sheet conformation.

Laser-light scattering technique enabled us to investigate the state of intermolecular interactions between peptide molecules, disclosing the existence of high molecular weight aggregates of PrP 106-126, detectable immediately after dissolution of the peptide in phosphate buffer, both at pH 5.0 and 7.0.

The behaviour of the peptide PrP 106-126 in deionized water, where smaller aggregates are formed, suggests that some short-range repulsion among the charged groups, not screened by the lower ionic strength solvent, acts locally, preventing large scale aggregation of the molecules.

Laser light scattering data also suggest that the aggregates of PrP 106-126 are organized as fibrils ruling out the possibility that they consist of single extended strips; since if this were so, the different aggregates would become entangled at high peptide concentrations, giving rise to an interactive behaviour, which was not observed in our experiments.

The overall data suggest an hypothesis about the organization of these high molecular weight aggregates: they are most likely organized as lamellar structure. This might be consistent with the characteristic amphiphilic pattern of the peptide molecule, that may support a ß-sheet organization of the fibrils, more evident at pH 5.0; this hypothesis may also explain the impossibility of scrambled PrP 106-126 to form aggregates as a direct consequence of the random distribution of hydrophobic and hydrophilic residues in its molecule. Finally, the partial resistance to protease digestion displayed by the peptide PrP 106-126 indicate that the conformational characteristics and the aggregation behaviour of this fragment might be related to the molecular properties of the protease-resistant PrP^{Sc}.

Acknowledgments

This paper was generously supported by the Fondazione Angelo and Angela Valenti, Milan, Italy, by CNR (National Research Council) Rome, Italy "Convenzione di Psicofarmacologia" and by the Italian Ministry of Health, Department of Social Services (OB and FT).We are grateful to Professor Mario Corti and Dr. Enrico Garattini for their constructive critical reading of the manuscript.

BEST AVAILABLE COPY

References

1. Prusiner, S.B. (1992) Biochemistry 49, 12277-12288.
2. Basler, K., Oesch, B., Scott, M., Westaway, D., Walchli, M., Groth, D.F., McKinley, M.P., Prusiner, S.B. and Weissman, C. (1986) Cell 46, 417-428.
3. Borchelt, D.R., Scott, M., Taraboulos, A., Stahl, N. and Prusiner, S.B. (1990) J. Cell Biol. 110, 743-752.
4. Borchelt, D.R., Taraboulos, A. and Prusiner, S.B. (1992) J. Biol. Chem. 267, 6188-6199.
5. Bolton, D.C., McKinley, M.P. and Prusiner, S.B. (1982) Science 218, 1309-1311. Sci. USA 89, 9349-9353.
6. Prusiner, S.B., Bolton, D.C., Groth, D.F., Bowman, K.A., Cochran, S.P. and McKinley, M.P. (1982) Biochemistry 21, 6942-6950.
7. McKinley, M.P., Bolton, D.C. and Prusiner, S.B. (1983) Cell 35, 57-62.
8. Prusiner, S.B., McKinley, M.P., Bowman, K.A., Bolton, D.C., Bendheim, P.E., Grooth, D.F. and Glenner, G.G. (1983) Cell 35, 349-358.
9. Prusiner, S.B., Grooth, D.F., Bolton, D.C., Kent, S.B. and Hood, L.E. (1984) Cell 38, 127-134.
10. Gasset, M., Baldwin, M.A., Fletterick, R.J. and Prusiner, S.B. (1993) Proc. Natl. Acad. Sci. USA 90, 1-5.
11. Gerstmann, J., Sträussler, E. and Scheinker, I. (1936) Z.ges Neurol. Psychiatr. 154, 736-762.
12. Farlow, M.R., Tagliavini, F., Bugiani, O. and Ghetti, B. (1991) In "Handbook of Clinical Neurology: Hereditary Neuropathies and Spinocerebellar Atrophies" (de Jong, J.M.B.V., Ed), Vol.60, pp.619-633, Elsevier, Amsterdam.
13. Ghetti, B., Tagliavini, F., Masters, C.L., Beyreuther, K., Giaccone, G., Verga, L., Farlow, M.R., Conneally, P.M., Dlouhy, S.R., Azzarelli, B. and Bugiani, O. (1989) Neurology 39, 1453-1461.
14. Giaccone, G., Verga L., Bugiani, O., Frangione B., Serban, D., Prusiner, S.B., Farlow, M.R., Ghetti, B. and Tagliavini, F. (1992) Proc. Natl. Acad.Sci. USA 89, 9349-9353.
15. Tagliavini, F., Prelli, F., Ghiso, J., Bugiani, O., Serban, D., Prusiner, S.B., Farlow, M.R., Ghetti, B. and Frangione, B. (1991) EMBO J. 10, 513-519.
16. Tagliavini, F., Prelli, F., Verga L., Giaccone, G., Salmona, M., Passerini, F., Wisniewski, T., Ghetti, B., Bugiani, O. and Frangione, B. (1992) Soc. Neurosci. Abstr. 18, 520.13.
17. Tagliavini, F., Prelli, F., Verga, L., Giaccone, G., Sarma, R., Gorevic, P., Ghetti, B., Ghibaudi, E., Passerini, F., Forloni, G., Salmona, M., Bugiani, O. and Frangione, B. (1993) Proc. Natl. Acad. Sci. USA, in press.
18. Forloni, G., Angeretti, N., Chiesa R., Monzani, E., Salmona, M., Bugiani, O. and Tagliavini, F. (1993) Nature 362, 543-546.
19. Curtis Johnson ,W. Jr. (1988) Ann. Rev. Biophys. Biophys. Chem. 17, 145-166.
20. Corti, M. (1985) In "Physics of Amphiphiles: Micelles, Vesicles and Microemulsions", (V. Degiorgio and M. Corti, Eds.) pp. 121-151, North Holland, Amsterdam.
21. Degiorgio, V., Corti, M. and Minero, C. (1984) Nuovo Cimento D3, 44-61.
22. Corti, M., and Degiorgio, V. (1981) J. Phys. Chem. 85, 711-717.
23. Gasset, M., Baldwin, M.A., Lloyd D.H., Gabriel, J.M., Holtzman, D.M., Cohen, F., Fletterick, R. and Prusiner, S.B. (1992) Proc. Natl. Acad. Sci. USA 89, 10940-10944.

BEST AVAILABLE COPY

roc. Natl. Acad. Sci. USA
ol. 92, pp. 11899–11903, December 1995
Cell Biology

Localization, trafficking, and temperature-dependence of the *Aequorea* green fluorescent protein in cultured vertebrate cells

(intracellular reporter/fusion protein/dexamethasone/transcription)

HIDESATO OGAWA*, SATOSHI INOUYE†‡, FREDERICK I. TSUJI‡, KUNIO YASUDA*, AND KAZUHIKO UMESONO*

Graduate School of Biological Sciences, Nara Institute of Science and Technology, 8916-5 Takayama, Ikoma, Nara 630-01, Japan; †Yokohama Research Center, hisso Corp., Kanazawa-ku, Yokohama, Kanagawa 236, Japan; and ‡Marine Biology Research Division 0202, Scripps Institution of Oceanography, niversity of California at San Diego, La Jolla, CA 92093

Communicated by Martin D. Kamen, Montecito, CA, September 11, 1995 (received for review July 7, 1995)

ABSTRACT **The localization, trafficking, and fluorescence of *Aequorea* green fluorescent protein (GFP) in cultured ertebrate cells transiently transfected with GFP cDNA were tudied. Fluorescence of GFP in UV light was found to be strongest when cells were incubated at 30°C but was barely visible at an incubation temperature of 37°C. COS-1 cells, rimary chicken embryonic retina cells, and carp epithelial ells were fluorescently labeled under these conditions. GFP was distributed uniformly throughout the cytoplasm and nucleus independent of cell type examined. When GFP was used to PML protooncogene product, fluorescence was detected in a unique nuclear organelle pattern indistinguishable from that of PML protein, showing the potential use of GFP as a fluorescent tag. To analyze both function and intracellular trafficking of proteins fused to GFP, a GFP–human glucocorticoid receptor fusion construct was prepared. The GFP–human glucocorticoid receptor efficiently transactivated the mouse mammary tumor virus promoter in response to dexamethasone at 30°C but not at 37°C, indicating that temperature is important, even for function of the GFP fusion protein. The dexamethasone-induced translocation of GFP–human glucocorticoid receptor from cytoplasm to nucleus was omplete within 15 min; the translocation could be monitored a single living cell in real time.**

The green fluorescent protein (GFP) from the jellyfish *Aequorea victoria* possesses the property of accepting energy by radiation-free energy-transfer process from an excited-state blue fluorescent protein during the aequorin bioluminescence reaction to emit light in the green wavelength (1–3). The protein has an absorbance maximum at 400 nm, with a minor peak at ≈480 nm, and its fluorescence-emission spectrum shows a sharp peak at 508 nm, with a shoulder at 545 nm (3). Chemical studies have shown that native GFP is a 27-kDa protein (4, 5) containing a cyclized tripeptide chromophore (5, 6). Molecular cloning and expression of GFP cDNA in *Escherichia coli* have led to amino acid-sequence determination of the protein and to the location of the chromophore in the primary structure (7, 8). GFP expressed in *E. coli* has also been shown to possess spectroscopic properties identical to the native protein (8–10). Recent studies have indicated that formation of the chromophore may depend on molecular oxygen or temperature (10–12). In addition, various mutant forms of GFP with different fluorescence excitation and emission spectra have been produced through mutagenesis (10, 12, 13).

The characteristic properties of GFP make this protein a good candidate for use as a molecular reporter to monitor patterns of protein localization, gene expression, and intracellular protein trafficking in living cells. Several such experiments have been recently carried out in *Caenorhabditis elegans* (8, 14), *Drosophila melanogaster* (15), and a mammalian cell line (16). In this study, expression constructs encoding GFP alone and GFP fused to other regulatory proteins were prepared, and expression of the proteins was analyzed in vertebrate cell lines and primary embryonic cells. Results show that the incubation temperature of GFP-expressing cells critically affects the intensity of green fluorescence and that under optimal conditions intracellular localization and trafficking of proteins can be visualized in a single living cell by fluorescence microscopy.

MATERIALS AND METHODS

Plasmids. The cloned GFP cDNA (pHis-AGP) (9) was used as a template for PCR to include a Kozak's translation start site and a 3′ linker sequence. The amplified DNA fragment was ligated to pCMX (17), generating pCMX-GFP. To prepare pCMX-PML–GFP, a *Mlu* I site was added to the 5′ end of the GFP coding sequence through PCR and inserted at the *Mlu* I site in pCMX-epi-PML (18). The human glucocorticoid receptor (hGR) cDNA, with a truncated 5′ coding region, was isolated from pGR107 (19) by *Sal* I–*Bam*HI digestion and ligated at the cognate sites of pCMX-GFP. The expression vectors used have already been described: pCMX-βGAL (17) and for wild-type hGR (19), amino-terminal truncated hGR mutant (ΔAD) (20), and mouse mammary tumor virus promoter (MMTV)-LUC (firefly) (20).

Cell Culture and Transfection. COS-1 cells, chicken embryonic retina cells, and EPC (fish cell line derived from carp epithelial tumor) cells were maintained in Dulbecco's modified Eagle's medium, without phenol red, supplemented with 10% fetal bovine serum. For EPC cells, 25 mM Hepes was added to the culture medium. For dexamethasone treatment, resin/charcoal-stripped fetal bovine serum was used. Retina cells were dissected from 8-day-old chicken embryos and cultured as described (21). Transfection was done in 6-cm tissue culture dishes by the calcium phosphate-DNA precipitation method for 6–9 hr at 37°C using 6 μg of DNA (17). After washing the DNA precipitate, the following was carried out: (*i*) COS-1 cells were kept for 24 hr at 37°C and then incubated for 36 hr at a specified temperature, (*ii*) retina cells were incubated at 30°C for 48 hr, and (*iii*) EPC cells were maintained at 30°C, including the transfection period (22). For luciferase assays, indicated expression vectors (0.2 μg) were cotransfected with MMTV-LUC (0.5 μg) and pCMX-βGAL (0.2 μg), in 24-well tissue culture dishes. Luminescence was measured in triplicate samples by using a Berthold (Nashua, NH) luminometer, according to the manufacturer's directions.

Microscopy and Image Analysis. Before microscopic observation, the medium was removed and replaced with phosphate-buffered saline buffer. The cells were examined by using a

e publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: GFP, green fluorescent protein; hGR, human glucocorticoid receptor; DAPI, 4′,6-diamidino-2-phenylindole; MMTV, mouse mammary tumor virus promoter.

BEST AVAILABLE COPY

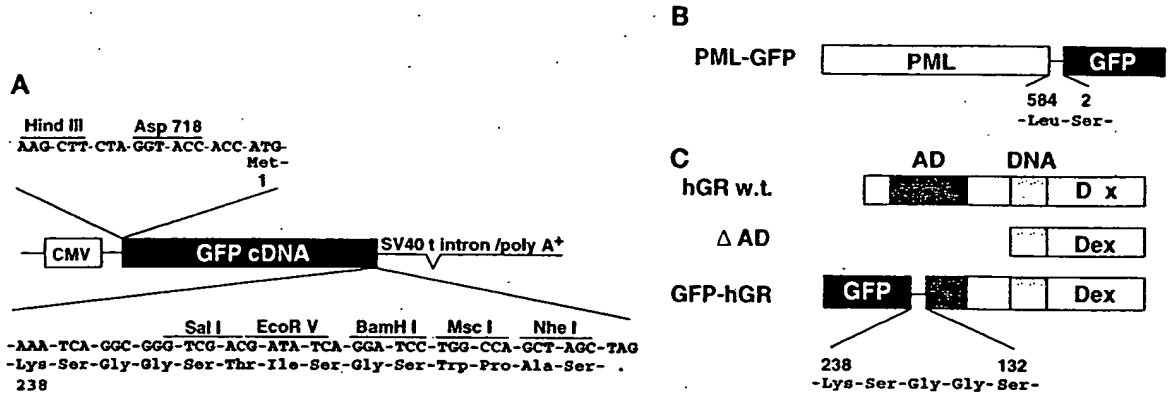


FIG. 1. Construction of GFP expression plasmids and structure of GFP fusion proteins. (*A*) Schematic structure of pCMX-GFP, encoding full-length GFP for expression in vertebrate cells. GFP cDNA is transcribed from human cytomegalovirus (CMV) immediate early promoter followed by a simian virus 40 small t intron and poly(A)$^+$ signal. The 5′ end of the cDNA was modified to encode a Kozak's consensus sequence (ACCACC-ATG). To the carboxyl-terminal lysine codon of the GFP cDNA, 13 amino acid residues were added to accommodate several restriction enzyme sites. (*B*) Structure of PML–GFP fusion protein. Full-length GFP was fused at the carboxyl terminus of PML protein through *Mlu* I sites in the cDNAs. In this construct, PML protein lacks the last tryptophan residue. (*C*) Structure of full-length hGR (hGR w.t.), amino-terminal truncated hGR (ΔAD), and GFP–hGR fusion proteins. A synthetic initiation codon was placed in front of the DNA-binding domain (DNA) of hGR to generate ΔAD protein, which lacks the amino-terminal transactivation domain (AD) (20). In GFP–hGR, GFP was fused amino-terminal to an hGR that lacks the first 131 amino acid residues. Dex, dexamethasone.

Zeiss Axiophot microscope, equipped with a fluorescein isothiocyanate filter set for fluorescence detection and a ×40 or ×63 water-immersion objective. Where indicated, cells were fixed with 4% (vol/vol) paraformaldehyde on the culture dish. To visualize nuclei, cells were treated with 0.1% Triton X-100 and soaked in 4′,6-diamidino-2-phenylindole (DAPI) at 50 ng/ml in phosphate-buffered saline. Photographs were taken by using Kodak Echtachrome 400, and films were subsequently processed to enhance sensitivity.

RESULTS

Detection of GFP Fluorescence in Cells. Attempts were first made to visualize fluorescence in a single cell by overexpress-

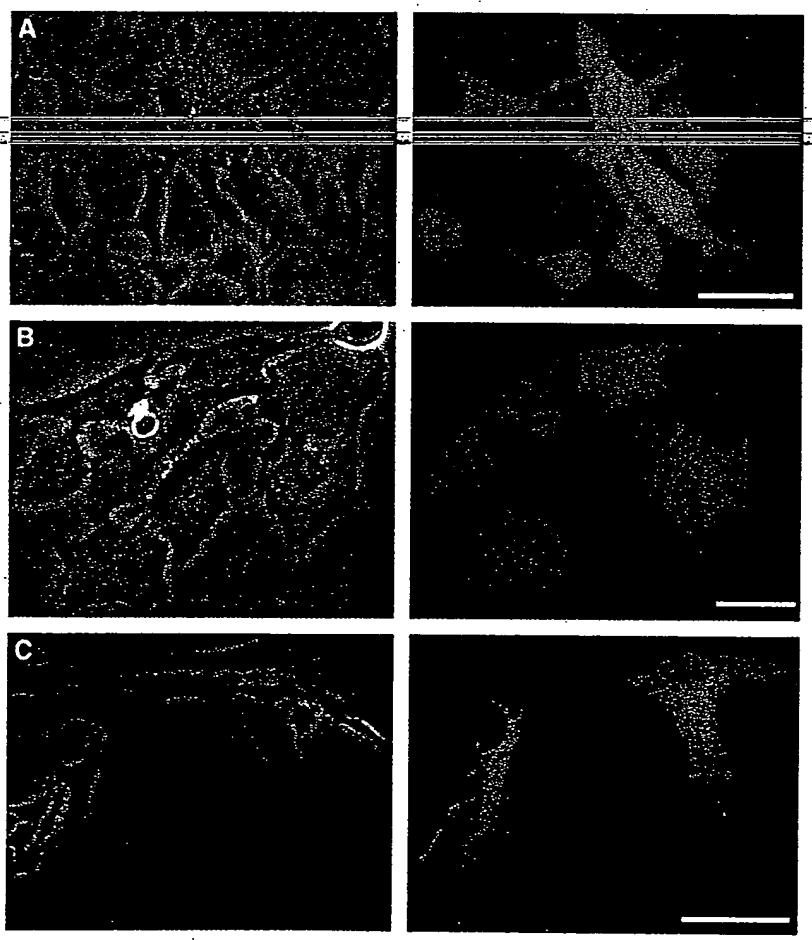


FIG. 2. Detection of fluorescence in cells expressing GFP. (*A*) COS-1 cells. (*B*) Embryonic chicken retina cells. (*C*) Carp epithelial EPC cells. (*Left*) Phase-contrast microscopy of cells. (*Right*) Fluorescence microscopy of same cells. (Bar = 25 μm.)

BEST AVAILABLE COPY

ing GFP in *E. coli* carrying pHis-AGP at 37°C (9). Strong green fluorescence in single cells was readily detectable under the microscope (data not shown). Subsequently, a eukaryotic expression construct of GFP, pCMX-GFP, was prepared and used for transient transfection of COS-1 cells. Fig. 1 shows a schematic diagram of GFP expression plasmids and structures of GFP fusion proteins. To monitor transfection efficiency, an expression plasmid for *E. coli* β-galactosidase (pCMX-βGAL) was cotransfected. The results showed that at 10% transfection efficiency, as judged by *in situ* staining of β-galactosidase activity, fluorescent cells were barely detectable. Similar results were obtained with CV-1 cells.

The low fluorescence intensity suggested the possibility that temperature may be a factor because *Aequorea victoria* lives in cold water (Friday Harbor, WA) and successful expression of GFP has been reported in nematodes and flies, both of which are usually maintained at 20–25°C (8, 14, 15). Experiments with temperature shifts showed that to detect fluorescence it was important to shift the incubation temperature after transfection from 37°C to 30°C for at least 4 hr before observing fluorescence (Fig. 2*A*). The strongest intensity was obtained when cells were kept at 37°C for 24 hr after removing the DNA precipitate and then incubating cells at 30°C for another 36–48 hr.

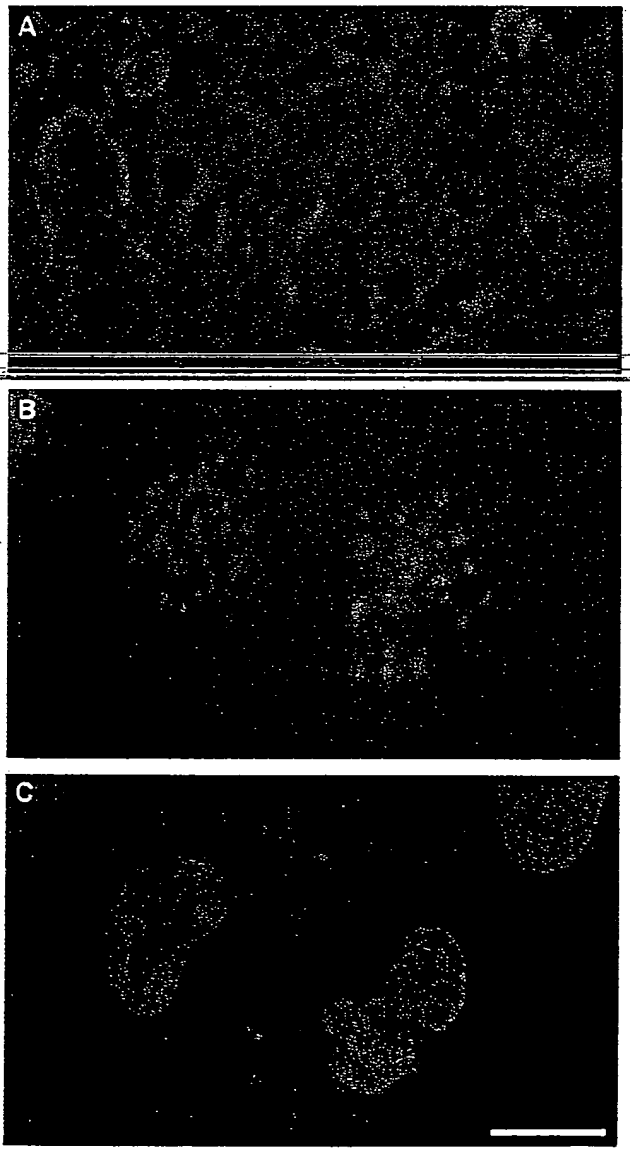


FIG. 3. Localization of fluorescently tagged PML protein in COS-1 cells. Same cells are viewed by phase contrast microscopy (*A*), fluorescence microscopy (*B*), and DAPI staining (*C*). (Bar = 10 μm.)

Similar temperature-shift experiments were done with embryonic cells in primary culture labeled with GFP. Retina cells, dissected from 8-day-old chicken embryos and maintained in culture, were transfected with pCMX-GFP at 37°C, followed by incubation at 30°C. Very good fluorescence was observed under these conditions (Fig. 2*B*), indicating that GFP can be used to label vertebrate embryonic cells, as well as immortalized cells.

An unavoidable problem encountered with mammalian and avian cells when the incubation temperature was lowered from 37°C to 30°C was a slowing in the metabolic rate. For comparison, the fish cell line EPC, derived from carp epithelial tumor and which proliferates well at 30°C (22), was transfected with pCMX-GFP and incubated at 30°C. The number of intensely fluorescent cells greatly increased over the number found with homoiothermic cells (Fig. 2*C*). The conclusions to be drawn from these observations are that temperature is an important parameter in detection of *in vivo* GFP fluorescence and that the promoter is highly active in driving transcription of cDNA for GFP.

Intracellular Localization of GFP. Fig. 2 shows that GFP appears uniformly distributed throughout the cytoplasm and nucleus. This pattern was not due to the additional 13-amino acid residues attached to the carboxyl terminus of the pCMX-GFP-encoded protein (Fig. 1*A*) because an identical pattern was observed with intact unmodified GFP (data not shown). To visualize other proteins, a host protein with an easily recognizable pattern of localization was selected. The *PML* protooncogene has been identified as a partner that undergoes reciprocal chromosomal recombination with the retinoic acid receptor α gene in acute promyelocytic leukemia (23–26). The PML protein localizes in punctate nuclear organelles called PML oncogenic domains or nuclear bodies (18, 27, 28). The GFP was fused to the carboxyl terminus of the PML protein (Fig. 1*B*), and expression plasmid pCMX-PML–GFP was used to transfect COS-1 cells. Upon incubation as described, uniform distribution of GFP fluorescence was no longer seen; instead a punctate nuclear fluorescence pattern, reminiscent of the PML oncogenic domains in fixed cells, was visualized

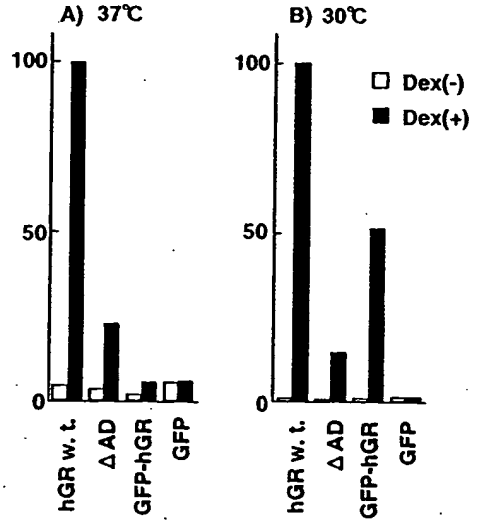


FIG. 4. Transcriptional enhancement by hGR mutants in COS-1 cells. Dexamethasone (Dex)-inducible reporter MMTV-LUC was cotransfected with expression plasmids encoding full-length hGR (hGR w.t.), amino-terminal truncated mutant ΔAD, GFP–hGR fusion, or GFP alone. Two different temperatures were tested: *A*, 37°C; *B*, 30°C. Maximum induction obtained with hGR w.t. by challenge with 1 μM Dex was taken as 100, and the relative reporter luciferase activities were plotted. All values are averages of triplicate experiments.

BEST AVAILABLE COPY

(Fig. 3*B*). A single fluorescent spot is called a "nuclear body," and the two clusters represent two cell nuclei. Most fluorescent spots were nuclear, but occasionally a few spots could be detected in the cytoplasm. This result showed that GFP can be used as an efficient tag to trace distribution of the host protein to small nuclear organelles such as PML oncogenic domains in living cells.

Properties of GFP–hGR Fusion Proteins in COS-1 Cells. To carry out real-time imaging of protein trafficking in a single living cell, pCMX-GFP–hGR was prepared by fusing GFP to a mutant hGR containing intact DNA- and ligand-binding domains (Fig. 1*C*). The advantages of this fusion protein are as follows: (*i*) it contains the functional properties of a transcription factor that can be addressed through hormone-dependent transactivation of reporter genes such as MMTV-LUC (20) and (*ii*) hGR, which is cytoplasmic in the absence of hormone, translocates to the nucleus upon binding to hormone (for review, see ref. 29).

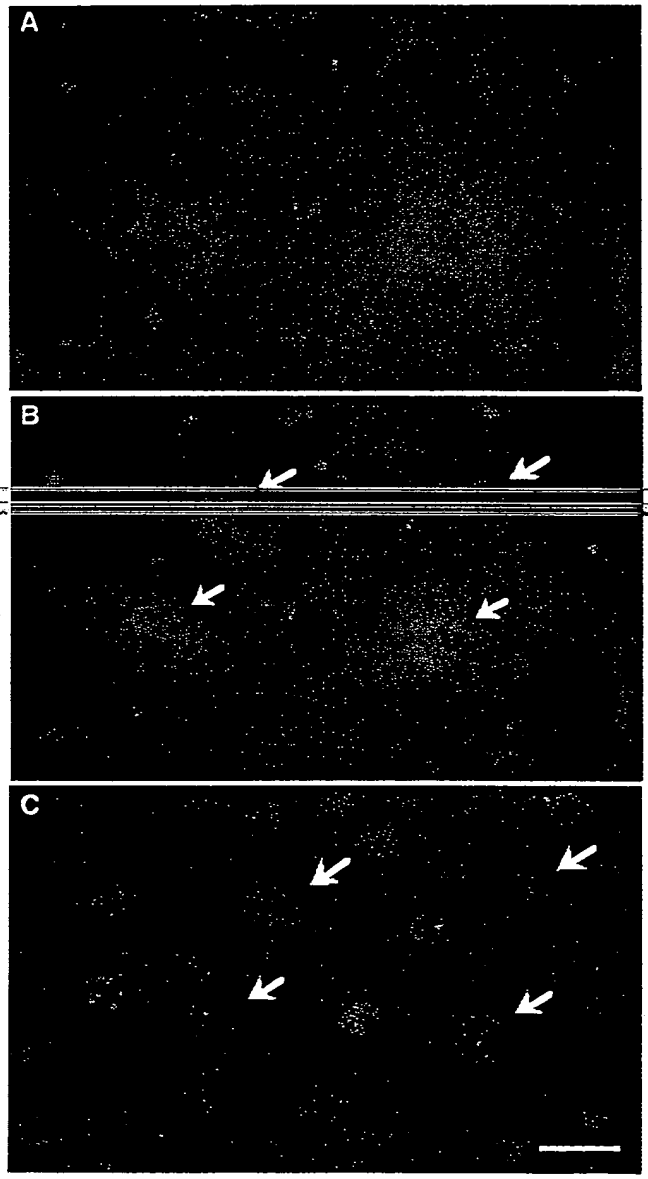


FIG. 5. Nuclear transfer of GFP–hGR fusion protein in COS-1 cells. (*A*) COS-1 cells were transfected with pCMX-GFP–hGR, and fluorescence was observed mainly in the cytoplasm. (*B*) Fluorescence seen 30 min after addition of 1 μM dexamethasone to COS-1 cells in *A*. Arrows indicate nuclear fluorescence from translocation of fusion protein in response to hormone. (*C*) Same field after DAPI staining. (Bar = 25 μM.)

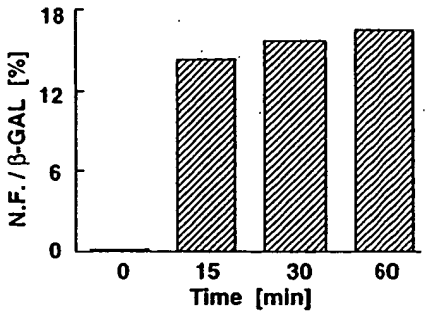


FIG. 6. Time course of increase in number of COS-1 cells with positive nuclear fluorescence after dexamethasone addition. Cells were transfected with pCMX-GFP–hGR and pCMX-βGAL, after which they were incubated for 0, 15, 30, and 60 min with 1 μM dexamethasone and then fixed. Number of cells showing nuclear fluorescence (N.F.) was determined, and cells were then analyzed for β-galactosidase (β-GAL) activity *in situ*. Percentages of fluorescent cells against β-galactosidase-positive cells at each time point were calculated and plotted.

With COS-1 cells, pCMX-GFP–hGR was cotransfected with MMTV-LUC, and transfection efficiency was determined by using cotransfected pCMX-βGAL. When cells were kept at 37°C, transfection with pCMX-GFP–hGR gave no transactivation, whereas wild-type hGR and truncated hGR lacking the amino-terminal transactivation domain (ΔAD) induced reporter luciferase activity in response to dexamethasone, a synthetic glucocorticoid (Fig. 4*A*). However, efficient transactivation by GFP–hGR was seen when the incubation temperature was shifted to 30°C. This result showed that GFP can be fused to a transcription factor without affecting its function of either binding DNA or interacting with the transcription machinery (Fig. 4*B*).

Further evidence of GFP localization and nuclear translocation was obtained as follows. Without dexamethasone, COS-1 cells transfected with pCMX-GFP–hGR showed localization of the fusion protein, as evidenced by green fluorescence, primarily in the cytoplasm (Fig. 5*A*). When dexamethasone (1 μM) was added, followed by incubation at 30°C for 30 min, fluorescence shifted into the nucleus (Fig. 5*B*).

To estimate the rate of GFP–hGR nuclear transfer induced by dexamethasone, COS-1 cells were cotransfected with pCMX-GFP–hGR (5 μg) and pCMX-βGAL (1 μg), and after visualization of cytoplasmic fluorescence, cells were challenged with hormone for 0, 15, 30, and 60 min at 37°C, followed by paraformaldehyde fixation to stop further protein transfer. The number of cells with nuclear fluorescence was determined, and percentages were calculated from the number of β-galactosidase-positive cells on the same plate. Fig. 6 shows that nuclear transfer was almost complete within 15 min. Further, direct observation under a microscope revealed that some cells completed their nuclear transfer of GFP–hGR within 5 min.

DISCUSSION

Since the report of Chalfie *et al.* (8), there has been wide interest in the application of GFP to solving biological problems. In this study, we have concentrated on GFP labeling in vertebrate cell systems. Our results show that incubation temperature is important in the detection of GFP fluorescence in mammalian, avian, and fish cells. Further, functional activity of fusion protein GFP–hGR also appears to depend on temperature. The underlying cause of this temperature dependence is still unknown, but it may be associated with the folding and/or redox state of GFP (11) within the cell. A number of factors appear to be involved because (*i*) fluorescence of highly purified recombinant GFP (11) dissolved in 100 mM ammo-

BEST AVAILABLE

nium bicarbonate, pH 8.0, did not show the same temperature dependence as the protein *in vivo*, (*ii*) incubation time required at 30°C was relatively short (4 hr), (*iii*) fluorescence was greater when a 24-hr incubation period at 37°C preceded the temperature shift, presumably due to GFP accumulation, and (*iv*) once the cells began to fluoresce at 30°C, they remained fluorescent when the temperature was shifted back to 37°C (at least for an additional 48 hr).

These results suggest various possibilities for using GFP as an intracellular reporter—for example, in labeling chicken embryos by retroviral gene transfer and shifting the temperature down during incubation. Marshall *et al.* (16) have recently reported the detection of GFP targeted to membrane in a human cell line after transient transfection with an expression construct. Even though the gfp10 GFP (7) used by Marshall *et al.* (16) differs at seven amino acid positions from our GFP (9), their protein has also been found to exhibit temperature sensitivity when expressed in yeast cells (K. Kohno, personal communication). Both the work by Marshall *et al.* (16) and our work clearly indicate that GFP may be used to visualize intracellular components within a vertebrate cell. Visualization of intracellular protein trafficking as seen with GFP–hGR should also allow experiments to be designed to address important biological questions. For example, immediate applications would be the study of nuclear translocation of other transcription factors (e.g., NF-κB, aryl hydrocarbon receptors, sterol regulatory element-binding protein 1, and STAT proteins), synaptic vesicle transfer, exocytosis, stress-induced cytoskeletal reorganization, and protein secretion.

We are indebted to Dr. R. T. Yu for critical reading of the manuscript and members of the Yasuda Lab for valuable discussions during the study. We thank Dr. R. M. Evans for hGR plasmids, Dr. A. Kakizuka for pCMX-epi-PML, Drs. H. Toyohara and M. Kinoshita for EPC cells, and Dr. K. Kohno for valuable discussions. This work was supported in part by grants from the Japan Ministry of Education, Science and Culture and National Science Foundation Grant MCB-9104684.

1. Johnson, F. H., Shimomura, O., Saiga, Y., Gershman, L. C., Reynolds, G. T. & Waters, J. R. (1962) *J. Cell. Comp. Physiol.* **60**, 85–103.
2. Morin, J. G. & Hastings, J. W. (1971) *J. Cell. Physiol.* **77**, 313–318.
3. Morise, H., Shimomura, O., Johnson, F. H. & Winant, J. (1974) *Biochemistry* **13**, 2656–2662.
4. Prendergast, F. G. & Mann, K. G. (1978) *Biochemistry* **17**, 3448–3453.
5. Shimomura, O. (1979) *FEBS Lett.* **104**, 220–222.
6. Cody, C. W., Prasher, D. C., Westler, W. M., Prendergast, F. G. & Ward, W. W. (1993) *Biochemistry* **32**, 1212–1218.
7. Prasher, D. C., Eckenrode, V. K., Ward, W. W., Prendergast, F. G. & Cormier, M. J. (1992) *Gene* **111**, 229–233.
8. Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W. & Prasher, D. C. (1994) *Science* **263**, 802–805.
9. Inouye, S. & Tsuji, F. I. (1994) *FEBS Lett.* **341**, 277–280.
10. Heim, R., Prasher, D. C. & Tsien, R. Y. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 12501–12504.
11. Inouye, S. & Tsuji, F. I. (1994) *FEBS Lett.* **351**, 211–214.
12. Heim, R., Cubitt, A. B. & Tsien, R. Y. (1995) *Nature (London)* **373**, 663–664.
13. Delagrave, S., Hawtin, R. E., Silva, C. M., Yang, M. M. & Youvan, D. C. (1995) *Bio/Technology* **13**, 151–154.
14. Sengupta, P., Colbert, H. A. & Bargmann, C. I. (1994) *Cell* **79**, 971–980.
15. Wang, S. & Hazelrigg, T. (1994) *Nature (London)* **369**, 400–403.
16. Marshall, J., Molloy, R., Moss, G. W. J., Howe, J. R. & Hughes, T. E. (1995) *Neuron* **14**, 211–215.
17. Umesono, K., Murakami, K. K., Thompson, C. C. & Evans, R. M. (1991) *Cell* **65**, 1255–1266.
18. Dyck, J. A., Maul, G. G., Miller, W. H., Jr., Chen, J. D., Kakizuka, A. & Evans, R. M. (1994) *Cell* **76**, 333–343.
19. Hollenberg, S. M., Weinberger, C., Ong, E. S., Cerelli, G., Oro, A., Lebo, R., Thompson, E. B., Rosenfeld, M. G. & Evans, R. M. (1985) *Nature (London)* **318**, 635–641.
20. Hollenberg, S. M. & Evans, R. M. (1988) *Cell* **55**, 899–906.
21. Okada, T. S., Yasuda, K., Araki, M. & Eguchi, G. (1979) *Dev. Biol.* **68**, 600–617.
22. Kamei, Y., Yoshimizu, M. & Kimura, T. (1987) *Fish Pathol.* **22**, 147–152.
23. Kakizuka, A., Miller, W. H., Jr., Umesono, K., Warrell, R. P., Jr., Frankel, S. R., Murty, V. V. V. S., Dmitrovsky, E. & Evans, R. M. (1991) *Cell* **66**, 663–674.
24. de Thé, H., Lavau, C., Marchio, A., Chomienne, C., Degos, L. & Dejean, A. (1991) *Cell* **66**, 675–684.
25. Goddard, A. D., Borrow, J., Freemont, P. S. & Solomon, E. (1991) *Science* **542**, 1371–1374.
26. Kastner, P., Perez, A., Lutz, Y., Rochette-Egly, C., Gaub, M.-P., Durand, B., Lanotte, M., Berger, R. & Chambon, P. (1992) *EMBO J.* **11**, 629–642.
27. Weis, K., Rambaud, S., Lavau, C., Jansen, J., Carvalho, T., Carmo-Fonseca, M., Lamond, A. & Dejean, A. (1994) *Cell* **76**, 345–356.
28. Koken, M. H. M., Puvion-Dutilleul, F., Guillemin, M. C., Viron, A., Linares-Cruz, G., Stuurman, N., de Jong, L., Szostecki, C., Calvo, F., Chomienne, C., Degos, L., Puvion, E. & de Thé, H. (1994) *EMBO J.* **13**, 1073–1083.
29. Akner, G., Wikström, A.-C. & Gustafsson, J.-Å. (1995) *J. Steroid Biochem. Mol. Biol.* **52**, 1–16.

BEST AVAILABLE COPY

The EMBO Journal vol.12 no.9 pp.3459 - 3466, 1993

Chimeric r tinoic acid/thyroid hormone receptors implicate RAR-α1 as mediating growth inhibition by retinoic acid

Johannes G.Schilthuis, Alexander A.F.Gann and Jeremy P.Brockes

Ludwig Institute for Cancer Research, and Department of Biochemistry and Molecular Biology, University College London, 91 Riding House Street, London W1P 8BT, UK

Communicated by J.Brockes

Retinoic acid (RA) affects the growth and differentiation of cells in culture, usually to decrease the growth rate. In amphibian limb regeneration RA has the remarkable ability to affect pattern formation by changing positional identity, but its initial action on the limb is to inhibit division of the blastemal progenitor cells. Newt limb blastemal cells also show this inhibition in culture. In order to investigate the role of different RA receptors (RARs) in the RA response, the hormone binding domain of the newt RARs α1 and δ1 was replaced with the corresponding region from the *Xenopus* thyroid hormone receptor-α (TR-α). In COS cells transfected with each of the chimeras, transcription was activated after exposure to thyroid hormone (T3). Their profile of activity on three different response elements was indicative of RAR specificity and not TR specificity. After transfection of cultured newt blastemal cells with a DNA particle gun, the chimeras were equally active in stimulating T3-dependent transcription of two different synthetic reporter genes. Blastemal cells were transfected with chimeras or control plasmids along with a marker plasmid expressing β-galactosidase, exposed to RA or T3 and labelled with [^{3}H]thymidine followed by autoradiography. The α1 chimera gave T3-dependent inhibition of growth, comparable to the effect exerted by RA itself, whereas the δ1 chimera and control plasmids were inactive. The results imply that RAR-α1 mediates the effects of RA on blastemal cell growth.

Key words: cell growth/retinoic acid/retinoic acid response element/thyroid hormone

Introduction

Retinoic acid (RA) affects the proliferation and differentiation f a wide variety of cell types both in culture and *in vivo*. Although certain exceptions have been proposed (Ide and Aono, 1988; Paulsen *et al.*, 1988), its general action in culture is to decrease the growth rate and often to promote differentiation of precursor cells or established cell lines (Strickland and Mahdavi, 1978; Breitman *et al.*, 1980; Kim *et al.*, 1987). Its anti-neoplastic properties have been underlined recently by the therapeutic use in provoking differentiation of human leukaemic promyelocytes carrying the t(15;17) translocation (Huang *et al.*, 1988). There has also been much interest in its remarkable ability to respecify positional identity in developing and regenerating limbs (Brockes, 1989; Stocum, 1991; Tabin, 1991; Bryant and Gardiner, 1992). Limb regeneration in urodele amphibians such as the newt or axolotl proceeds by local formation of a blastema, a growth zone of mesenchymal progenitor cells which give rise to the regenerate (Wallace, 1981). The initial effect of RA on the blastema is to inhibit division (Maden, 1983), but subsequently an RA-treated blastema gives rise to extra structures that are indicative of a unidirectional change in axial specification (Maden, 1982; Stocum and Crawford, 1987). These diverse effects of RA are thought to be mediated by nuclear receptors of the steroid/thyroid superfamily which act as ligand-dependent transcription factors.

Studies in mouse and man have led to the identification of three genes coding for the retinoic acid receptors RAR-α, RAR-β and RAR-γ (Giguere *et al.*, 1987; Petkovich *et al.*, 1987; Benbrook *et al.*, 1988; Brand *et al.*, 1988; Zelent *et al.*, 1989). Sequence comparison of the receptors has shown that whereas the DNA-binding and ligand-binding domains are well conserved, there is considerable divergence in the NH_2-terminal A region. In addition each gene encodes multiple isoforms, of which the principal ones derive from alternative promoter usage and splicing at the NH_2-terminus (Kastner *et al.*, 1990; Leroy *et al.*, 1991; Zelent *et al.*, 1991). The distribution of the RARs has been analysed in detail, particularly during mouse development, by *in situ* hybridization with probes specific for each of the three RARs but not for individual isoforms (Dolle *et al.*, 1989, 1990; Ruberte *et al.*, 1990, 1991). These studies indicate that while RAR-α appears to be ubiquitous, RAR-β and RAR-γ show marked spatial and temporal regulation during embryogenesis. Furthermore, a study of transactivation by the various isoforms in transfection assays with different reporter genes has revealed clear differences in their activity that are dependent on the promoter and cell context (Nagpal *et al.*, 1992). Nonetheless there is little information that allows distinctions to be drawn about which isoform is responsible for mediating a particular physiological response to RA. A recent study with an RA-resistant subclone of human HL-60 leukaemic cells has shown that RAR-α, -β and -γ as well as RXR-α are all able to mediate granulocytic differentiation (Robertson *et al.*, 1992). In other experiments on HL-60, a synthetic RAR-α antagonist has been shown to counteract RA effects (Apfel *et al.*, 1992). It is clearly a challenge to identify the precise contributions of each isoform in different cell types, and to evaluate the possibility of functional redundancy.

In this report, we focus on the ability of RA to decrease the growth rate of cultured limb blastemal cells from the newt. The major isoform expressed in the newt limb and blastema is δ1 (Ragsdale *et al.*, 1989, 1992), which appears to be the urodele equivalent of γ1, although these receptors have diverged extensively in the A region. In addition, we have identified an α1 isoform which has high sequence

© Oxford University Press

BEST AVA

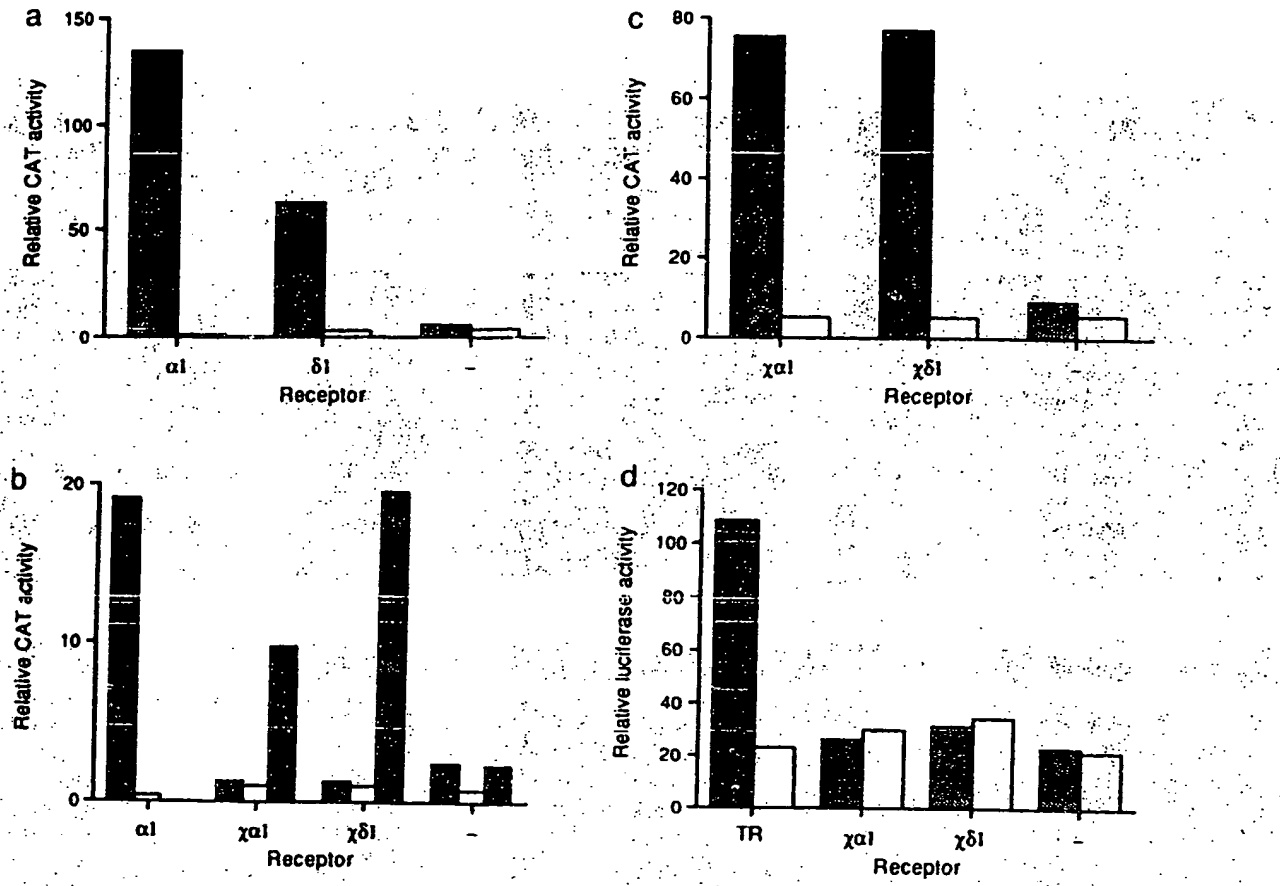


Fig. 1. (a) Activity of newt RARs in COS cells. COS cells were transfected with 5 μg expression plasmid for RARs, or vector control, 5 μg $(TRE3)_3$-tk-CAT and 5 μg EFβ-gal as a standard for transfection efficiency. The relative CAT activity is expressed in thousands of units after normalizing for β-galactosidase activity. Filled bars: cells treated with RA (10^{-7} M). Empty bars: untreated cells. The data shown in this and the subsequent figures are representative of that obtained in 3–5 experiments. (b) Activity of chimeric receptors on the palindromic TRE in COS cells. Cells were transfected with 5 μg expression construct for the chimera, or RAR-α1, or vector alone, 5 μg $(TRE3)_3$-tk-CAT and 5 μg EFβ-gal as above. After incubation in 10^{-7} M RA (filled bars), 5×10^{-7} M T3 (hatched bars) or no hormone (empty bars), the cells were extracted and the normalized CAT activity was determined. Note that the chimeras activate with T3 but not RA, while RAR-α1 activates with RA. (c) Activity of chimeric receptors on an RARE in COS cells. Cells were transfected with 5 μg receptor construct and 5 μg EFβ-gal as above, along with 5 μg RARE-tk-CAT. The normalized CAT activity was determined after incubation with T3 (hatched bars) or no hormone (empty bars). (d) Absence of activity of chimeric receptors on the direct repeat TRE. COS cells were transfected with 5 μg chimera, or TR-α expression vector, or vector alone, 5 μg EFβ-gal, 5 μg pMOMLV-tk-luc and 5 μg $(TRE3)_3$-tk-CAT. The normalized luciferase activity was determined after incubation with T3 (hatched bars) or no hormone (empty bars). As control, the CAT activity was measured to confirm that each chimera could activate another reporter in the same experiment (data not shown). The chimeras do not activate through the direct repeat TRE, but the TR-α control does so.

identity to its mammalian counterpart (Ragsdale *et al.*, 1989). We have replaced the RA binding domain of both newt isoforms with the corresponding region of the *Xenopus* thyroid hormone (T3) receptor-α (TR-α). The chimeras were comparably activated by T3 to stimulate expression of RA reporter genes but only the chimera of RAR-α1 inhibited growth in a T3-dependent fashion. This ability to mimic the quantitative effect of RA implicates the α1 receptor as a natural mediator of this response.

Results

Construction of chimeric retinoid/thyroid receptors

Chimeric receptors were constructed by replacing the E and F regions of the newt α1 and δ1 RARs with the corresponding region of the *Xenopus* TR-α. An *Eco*RI site is present in the δ1 sequence in the E region at a position 11 amino acids from the D/E boundary (Ragsdale *et al.*, 1989). This site lies upstream of the Ti and dimerization domains of the E region (Laudet *et al.*, 1992). Equivalent *Eco*RI sites were introduced into the RAR-α1 and TR-α genes by mutation, and the chimeric receptors, χα1 and χδ1, were assembled and introduced into an expression vector (see Figure 6). The construction did not introduce or change any residues at the junction, but resulted in direct apposition of the two sequences.

Activity of the chimeric receptors in COS cells

Figure 1a shows the activities of the parent newt RARs after transfection of expression constructs into COS cells along with a synthetic reporter gene carrying three copies of the palindromic T3 response element (TRE) of the mouse growth hormone promoter. This response element is activated by both RARs and TRs (Umesono *et al.*, 1988). A strong RA-dependent stimulation of reporter activity was observed for the newt receptors, as previously reported (Ragsdale *et al.*, 1989). The chimeric receptors were analysed in the same system along with an α1 control. As shown in Figure 1b, the chimeras were not activated by RA, in contrast to the α1 control, whereas they were strongly activated by

BEST AVAILABLE COPY

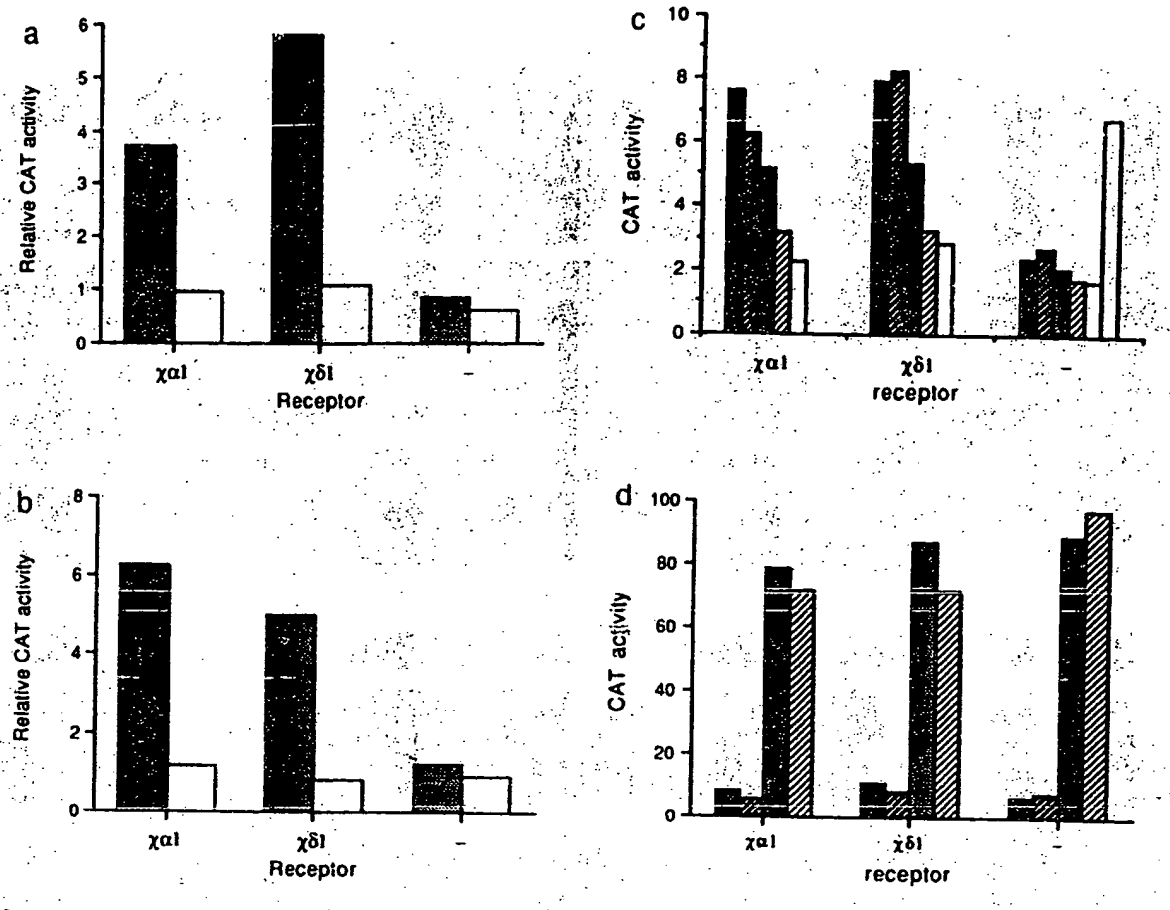


Fig. 2. (a) Activity of chimeric receptors on the palindromic TRE in cultured newt limb blastemal cells. Newt B1H1 cells were transfected using a particle gun with 2 μg chimeric receptor expression construct, or vector, 2 μg (TRE3)$_3$-tk-CAT and 2 μg EFβ-gal. The normalized CAT activity is shown for cells incubated in 5×10^{-7} M T3 (hatched bars) or without hormone (empty bars). (b) Activity of chimeric receptors on an RARE in newt limb cells. Newt cells were transfected with 2 μg receptor expression construct, or vector, 2 μg RARE-tk-CAT and 2 μg EFβ-gal. The normalized CAT activity is shown for cells incubated in T3 (hatched bars) or without hormone (empty bars). (c) Dose response for T3-induced stimulation of RARE by chimeras in newt cells. Newt cells were transfected as in (b) and CAT activity was determined after incubation in 0 (□), 10^{-9} (▨), 10^{-8} (▩), 10^{-7} (▨) or 10^{-6} (■) M T3. Also shown is the level of activation of the same reporter mediated by the endogenous RARs on treatment of the cells with 10^{-7} M RA (▩). (d) The chimeric receptors do not squelch transcription activated by an unrelated activator. Newt cells were transfected with 1 μg RGR expression construct, 1 μg GRE-tk-CAT reporter, 2 μg EFβ-gal and 2 μg chimera expression construct or vector. The normalized CAT activity is shown for cells incubated in 10^{-7} M RA (▩), 5×10^{-7} M T3 (▨), 10^{-7} M RA and 5×10^{-7} M T3 together (▨), or with neither RA nor T3 (■). In the presence of RA, reporter expression is activated by the RGR hybrid receptor; this activation is not significantly affected by either chimera.

5×10^{-7} M T3. A comparable stimulation by T3 was observed when the chimeras were assayed after cells were co-transfected with a reporter carrying the response element of the RAR-β2 promoter which is stimulated by RARs but not TRs (Figure 1c) (de The *et al.*, 1990; Sucov *et al.*, 1990). In contrast, the chimeras were not active on a reporter carrying the thyroid response element from the Moloney murine leukaemia virus (MoMLV) LTR, which is known to be specifically activated by TRs (Figure 1d) (Vivanco Ruiz *et al.*, 1991). As expected, this reporter was stimulated after co-transfection with an expression construct for the *Xenopus* TR-α (Figure 1d). Thus the activity of the chimeras in COS cells indicates that they retain the specificity for response elements that is characteristic of the RARs.

Activity of the chimeras in newt limb blastemal cells

The cultured newt cells used in these experiments were originally derived from a hind limb blastema (Ferretti and Brockes, 1988). They express blastemal mesenchyme markers such as reactivity with the monoclonal antibody 22/18 (Kintner and Brockes, 1985; Ferretti and Brockes, 1988) and antibodies to the K8/K18 cytokeratins (Ferretti *et al.*, 1989). The transfection efficiency for these cells is very low with conventional procedures, but using a DNA particle gun (Klein *et al.*, 1987; Yang *et al.*, 1990; Tang *et al.*, 1992), we routinely achieved frequencies up to ~10%, and this method was used for all of the experiments reported here. After transfection with reporters carrying the palindromic TRE (Figure 2a), or the retinoic acid response element (RARE) of the β2 promoter (Figure 2b), the chimeras showed similar T3-dependent stimulation of reporter activity. The concentration dependence of stimulation on the β2-RARE reporter in the range 10^{-6} to 10^{-9} M T3, a range covering that used in subsequent experiments (see below), is shown in Figure 2c. Also shown in this figure is that the fold-stimulation and magnitude of activation by the chimeras were comparable to those seen when the cells were stimulated with 10^{-7} M RA. This level of RA-dependent activation was even seen in cells expressing a chimeric receptor, demonstrating that the chimeras are not

BEST AVAILABLE COPY

expressed at a level that is sufficient to act as a dominant negative to the endogenous RARs (data not shown; Barettino *et al.*, 1993; Damm *et al.*, 1993).

In a separate experiment, the chimeras were shown not to squelch transcriptional activation by an unrelated activator. Expression from a reporter plasmid containing a glucocorticoid response element (GRE) upstream of the thymidine kinase (tk) promoter driving the chloramphenicol acetyltransferase (CAT) gene is activated in an RA dependent manner by an activator containing the DNA binding domain of the human glucocorticoid receptor and the RA binding domain of the human RAR-α. In newt cells, neither the α1 nor δ1 chimera, in the presence or absence of T3, affected transcriptional activation in this system (Figure 2d).

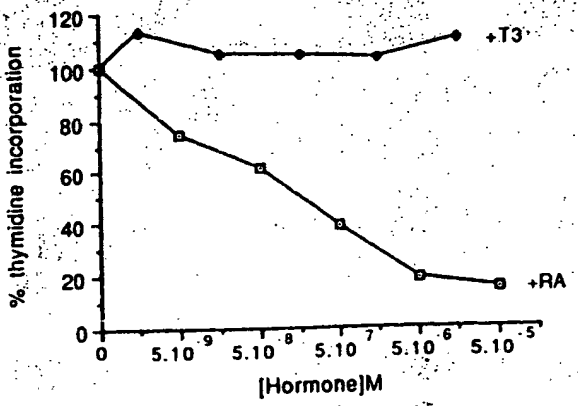


Fig. 3. Effect of T3 and RA on DNA synthesis in cultured newt cells. Newt cells were incubated with various concentrations of T3 (closed circles) or RA (open squares) for 72 h, followed by [^{3}H]thymidine for 14 h. The ^{3}H radioactivity incorporated into DNA was determined as described in Materials and methods, and expressed relative to the incorporation in cells incubated without hormone.

These observations indicate first that the chimeric receptors activate transcription in newt cells to an extent that is quantitatively similar to that produced by the endogenous RARs. Second, such activation is not limited by the availibility of other components, such as the RXRs. If this were the case the chimeras would have acted as dominant negatives to the endogenous RARs (Barettino *et al.*, 1993), yet they do not. Third, the chimeras do not interfere with transcription mediated by an unrelated activator, and hence do not exhibit squelching activity.

These findings encouraged us to analyse the functional consequence of chimera activation in newt cells.

Effect of chimera activation on growth of cultured blastemal cells

When cultured newt limb blastemal cells were exposed to various concentrations of RA for 3 days, there was a dose-dependent decrease in the growth rate as measured by incorporation of [^{3}H]thymidine (Figure 3), or by counting cell number (data not shown). If T3 was applied in the same range of concentrations to parallel cultures, no inhibitory effect was observed (Figure 3). No obvious change in cell differentiation was observed with either hormone. After transfection of newt limb cells with plasmid DNA using the DNA particle gun, transfectants continued to express plasmid

Fig. 4. Autoradiography of transfected newt cells. Cells were transfected with receptors and EFβ-gal as described in Figure 5, and incubated in T3 followed by [^{3}H]thymidine. After X-gal staining, the cells were coated with emulsion for autoradiography. One transfected (blue) cell has incorporated label and one has not. The background untransfected (white) cells show both positive and negative nuclei. The scale bar indicates 100 μm.

BEST AVAILABLE COPY

markers for at least 3 weeks in culture and began to incorporate [^{3}H]thymidine 7 days after transfection. It is therefore possible to determine the effect on the growth rate of activating chimeric receptors with T3.

Newt cells were transfected with chimeric or control constructions, exposed to varying concentrations of RA or T3, and incubated with [^{3}H]thymidine prior to autoradiography. In order to identify the transfected cells, all transfection mixtures contained plasmid DNA that expressed β-galactosidase. Co-transfection occurred in at least 85% of the recipients (see Materials and methods). The β-galactosidase was detected by X-gal staining, thus allowing the growth rate to be assessed by the proportion of transfected cells with silver grains over their nuclei, as illustrated in Figure 4.

The results from several experiments showed that when cells transfected with $\chi\alpha 1$ were activated with varying concentrations of T3, the decrease in growth rate was reproducibly similar to that observed when the same cell populations were treated with RA (Figure 5a). In contrast, when cells were transfected with $\chi\delta 1$ there was no effect on the growth rate in the equivalent range of T3 concentrations. It should be noted, as stated above, that the activity of the chimeras was comparable when assayed by transactivation of reporter genes. Cells transfected with either RAR-α1 or TR-α (the parent molecules of $\chi\alpha 1$) showed no T3-dependent inhibition of growth, while parallel transfections with $\chi\alpha 1$ gave the expected T3-dependent effect (Figure 5b).

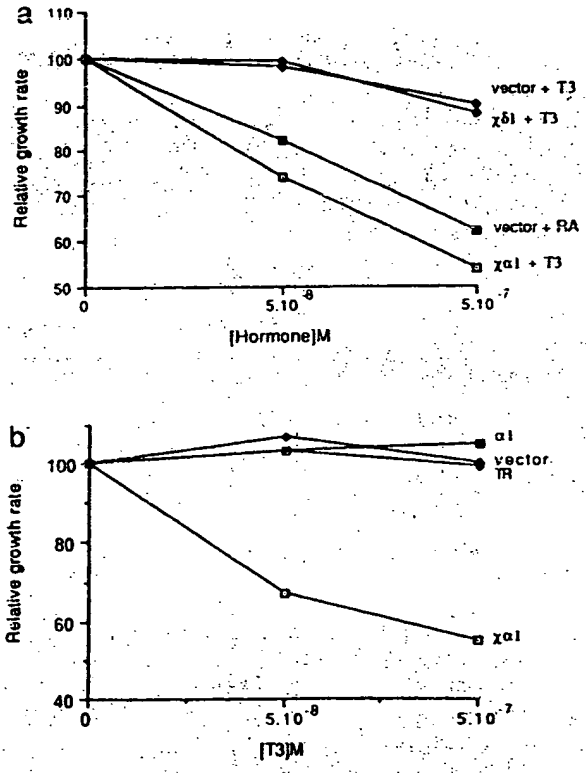


Fig. 5. Effect of chimera activation on newt limb cell growth. Cells were transfected with 2 μg chimera expression construct, or normal RAR-α1, or vector, and 2 μg EFβ-gal. After incubation with various concentrations of T3 or RA, the cultures were analysed as described in Materials and methods. The graphs show the relative growth rate of the transfected cells versus the concentration of hormone. In each case the graph shows results from a single experiment that is representative of data from at least three experiments. (a) Effect of T3 on growth of cells expressing $\chi\alpha 1$ or $\chi\delta 1$ as compared with the effect of RA or T3 on cells transfected with vector. Only $\chi\alpha 1$ is able to give a T3-dependent inhibition of cell growth comparable to that given by RA. (b) Effect of T3 on the growth of cells expressing vector, $\chi\alpha 1$, or the 'parental' receptors RAR-α1 and TR-α. Only $\chi\alpha 1$ gives the T3-dependent inhibition of growth. The total number of transfected cells with labelled nuclei counted for each category in all experiments was as follows: vector + T3, 1935; vector + RA, 807; $\chi\alpha 1$, 1736; $\chi\delta 1$, 542; $\alpha 1$, 2151; TR, 844.

Discussion

An advantage of the present approach is that it allows activation of one receptor isoform at a time in cells that express several different RARs or RXRs. The choice of the T3 binding domain to replace the RA binding domain was made for two reasons: T3 has no marked effect on limb regeneration (Hay, 1956) and, as the TR is a close relative of RARs in the nuclear receptor superfamily (Laudet *et al.*, 1992), it seemed possible that the T3 binding domain would function to activate transcription in the context of the RAR. Nonetheless it was important to determine if the hybrid receptors behave as RARs in respect of their transactivation properties. In the present case, each of the chimeras retained the specificity for response elements that is characteristic of the RAR as opposed to the TR. In addition the degree of transactivation was comparable to that observed with the parental RARs. It should be noted that we have not explored the effect of varying the precise contribution from the RAR and TR-α, and it might be important that the present constructions retain 11 amino acids from the E region of the RARs. Although the two hybrids described here behaved in an orderly fashion, we have found other cases where replacing the RA binding domain resulted in nonfunctional receptors (data not shown). In any event, the availability of functional chimeras allowed us to investigate the basis of an important effect of RA, that on cell growth in culture.

When the cultured blastemal cells were transfected using a particle gun, the cells continued to express transfected markers for at least 3 weeks. Transfected cells started to incorporate tritiated thymidine 1 week after transfection, and we have verified that the quantitative effects of RA on incorporation were comparable for untransfected cells and cells 7–10 days after transfection. This has allowed us to investigate chimera activity in the context of a 'transient' transfection assay. The most important result was that activation of the $\chi\alpha 1$ receptor with varying concentrations of T3 produced a decrease in the growth rate which reproducibly paralleled that obtained with RA, whereas control transfections with plasmid vector, RAR-α1 or TR-α had no effect. This strongly supports the view that this effect of RA is mediated by RARs, and more specifically is in agreement with properties attributed to RAR-α, though not to a specific isoform, by use of a putative α-selective antagonist (Apfel *et al.*, 1992). Since RXRs are not activated by T3 (Mangelsdorf *et al.*, 1990), it also demonstrates that ligand activation of RXRs is not necessary for this effect.

In contrast to $\chi\alpha 1$, the $\chi\delta 1$ construction gave no significant decrease in growth rate. When analysed for transactivation activity in newt cells, the two chimeras were the same, and it is likely, therefore, that the failure of $\chi\delta 1$ to affect the growth rate reflects a genuine difference in receptor function between RAR-δ1 and RAR-α1. We do not believe that the inhibiti n of growth by $\chi\alpha 1$ is the result of a non-specific

BEST AVAILABLE COPY

Somatic Cell and Molecular Genetics, Vol. 20, No. 3, 1994, pp. 153–162

Membrane-Bound Neomycin Phosphotransferase Confers Drug-Resistance in Mammalian Cells: A Marker f r High-Efficiency Targeting of Genes Encoding Secreted and Cell-Surface Proteins

William A. Mohler and Helen M. Blau

Department of Molecular Pharmacology, Stanford University, Stanford, California 94305-5332

Received 11 April 1994

Abstract—*An efficient method for inactivating genes is the use of silent selectable markers that are expressed only after homologous recombination into the active target gene. However, use of this approach for genes encoding secreted or membrane-anchored proteins may produce hybrid proteins comprising the N-terminal signal sequence from the target gene linked to the protein conferring drug resistance. Such chimeric enzymes will be secreted, precluding selection for drug resistance. To overcome this problem, we tested the possibility of anchoring in the membrane the cytoplasmic neomycin phosphotransferase (NPT). We constructed a fusion gene with a transmembrane domain connecting the N-terminal signal sequence of a membrane-targeted protein and the* neo *gene. Expression of this gene yielded G418-resistant colonies of C2C12 cells which contained assayable NPT activity. Comparison of enzyme activity in cell extract fractions verified that the active fusion protein was insoluble, presumably through localization to a membrane compartment. Transmembrane* neo *cassettes should serve as integration-activated markers capable of targeting genes encoding secreted or cell surface proteins.*

INTRODUCTION

When the expression of genes encoding proteins that confer drug resistance is made conditional upon homologous recombination, the efficiency of targeting active genes is very high, often exceeding one homologous event for every two selected clones (1–3). In such "expression-trap" targeting constructs, the selectable marker gene, which has neither an active promoter nor a translation start-codon in the vector, produces functional protein only after integration into a transcribed open reading frame. Homologous recombination of the vector into the target gene results in expression of a target/marker fusion protein in lieu of the wild-type gene product, and the outcome is a functional disruption of the endogenous gene. The high frequency of correctly targeted recombination following drug selection is likely due to the limited number of integration sites that can activate the selectable marker. Although genomic integration of a constituitively active marker at any site in any orientation can result in drug resistance, activation of a marker gene lacking a promoter and translation-start site requires insertion into an active gene, in the sense orientation, and within the actively translated reading frame.

In constructs that take full advantage of

0740-7750/94/0500-0153$07.00/0 © 1994 Plenum Publishing Corporation

BEST AVAILABLE COPY

Material may be protected by copyright law (Title 17, U.S. Code)

the selective power of the expression-trap approach, the codons for translation initiation and the *N*-terminus of the endogenous protein are required for the selectable marker gene to be expressed. Unfortunately, the structures of genes encoding most secreted or cell-surface proteins, such as hormones, growth factors, cytokines, adhesion molecules, and components of the extracellular matrix, currently make these genes untenable targets for homologous recombination via expression-trap selection. Short signal sequences that direct these proteins to the endoplasmic reticulum, or for subsequent secretion from the cell, reside in the extreme *N*-terminus of many such proteins (4). Because homologous recombination requires long regions of genomic DNA to be included in the vector, the prospects of deleting the signal sequence while retaining use of the endogenous ATG for translation of the marker are small. Thus target/marker fusion proteins created by expression-trap of these genes will retain an *N*-terminal signal sequence and are likely to be secreted, resulting in a drug-sensitive cell.

NPT expressed from the *neo* gene has previously been shown to function as a drug-selectable marker when produced as a fusion protein with various other protein domains (3, 5–7). Furthermore, NPT inserted into the cytoplasmic tail of a transmembrane protein has been used efficiently to select for homologous recombination into a target gene encoding a transmembrane receptor (8). However, in this case, the fusion protein expressed from the targeted gene retained all extracellular and transmembrane portions of the wild-type receptor, and mutant cells expressed undiminished levels of the native extracellular domain on their surfaces. Thus, although this approach is successful in disrupting signaling through the mutant receptor, it could not be used to target genes for secreted proteins or to eliminate expression of cell-surface proteins.

In this study we have designed and tested a construct that overcomes these problems and should allow expression-trap selection for homologous recombination in genes encoding secreted and cell surface proteins. We show here that like the cytoplasmic enzyme bacterial β-galactosidase (9), NPT can function when closely apposed to a hydrophobic membrane-spanning domain. We introduced a transmembrane (TM) domain *N*-terminal to the marker gene *neo*. When this tandem arrangement of domains is used to interrupt the extracellular sequence of a transmembrane protein, the TM domain appears to prevent secretion and anchor the protein in the membrane, retaining NPT activity within the cell and producing selective drug resistance. As this TM/*neo* cassette is as easily introduced into constructs as the native *neo* sequence, it should prove generally useful for expression-trap selection to alter or inactivate genes encoding secreted, cell surface, and extracellular matrix proteins.

MATERIALS AND METHODS

Plasmids and Oligonucleotides. The murine neural cell adhesion molecule (NCAM) cDNA clone, NCAM.H3, has been described (10); the sequence of the region used in our construct is as published (11). The cDNA for human insulin receptor (huIR) was provided by Dr. Richard Roth; the sequence of the region used in our construct is as published (12). Maps and sequence for pMC1*neo*-polyA (13) and pSV2*neo* (14) are published.

Primers were synthesized on a PCR-mate DNA synthesizer using 40-nmol columns and synthesis reagents from Applied Biosystems. Oligonucleotides were cleaved from the columns in ammonium hydroxide and deblocked for at least 6 h at 55°C. Ammonia was evaporated in open tubes at 55°C, sodium chloride was added to 0.5 M, and solutions were butanol-extracted twice before ethanol precipitation. Pelleted primers were resuspended in water and stored at

4°C. Primers were
reaction (PCR)
further purificati
as follows: (#1)
CAAACGACCC
TACGCTTCCG
(#3A) 5′-TTC
GAAGCGTAA
(#4A) 5′-AAC
GTCCCGTCAA
ATTTGACC
GATCTTCACG
5′-CGGCAAG
GCACT-3′.

Recombinant
PCR construction
in Figure 1. Seq
plates were ampl
DNA Thermal C
polymerase (Pe

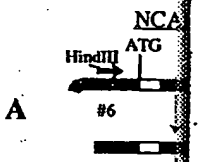


Fig. 1. Construction
product. Thin lines
Arrowheads show ex
translation start cod
individually amplified
produced by annealing
end-specific primers
huIR-TM (light grey)
sequence (dark grey;
relevant restriction si
strictly to scale.

Material may be protected by copyright law (Title 17, U.S. Code)

BEST AVAILABLE COPY

4°C. Primers were used for polymerase chain reaction (PCR) and sequencing without further purification. Primer sequences were as follows: (#1) 5'-GCGCGGATCCGAACAAACGACCC-3'; (#2) 5'-ACCGCTTTACGCTTCCGGATCGGCCATTGAA-3'; (#3A) 5'-TTCAATGGCCGATCCGGAAGCGTAAAGCGGTCCCAGC-3'; (#4A) 5'-AACGTGAAGATCTTCGACGTCCCGTCAAATATTGC-3'; (#5A) 5'-ATTTGACGGGACGTCGAAGATCTTCACGTTGACAGT-3'; and (#6) 5'-CGGCAAGCTTGCAAGGGGAAGGCACT-3'.

Recombinant PCR. The recombinant PCR construction scheme used is presented in Figure 1. Sequences from plasmid templates were amplified by standard PCR in a DNA Thermal Cycler using AmpliTaq DNA polymerase (Perkin Elmer/Cetus). The huIR-TM and *neo* sequences were then joined by recombinant PCR (described below), the product was gel isolated, and the joining reaction was used again to link NCAM with the huIR-TM/*neo* product to create the final recombinant PCR product, NTNeo.

Step A of Fig. 1 depicts the amplification of the component fragments used t create the fusion gene. PCR using primers specific to NCAM.H3 produced a 428-bp product that includes 70 bp of 5'-UTR, the translation start site (ATG in figure), and the first 115 codons of NCAM containing the signal sequence (white box). The upstream primer (#6) adds 6 bp and a new HindIII site to the 5-end of the product, while the downstream primer (#5A) adds a 15-bp tail of huIR-TM sequence and creates the in-frame linkage of NCAM and huIR-TM in

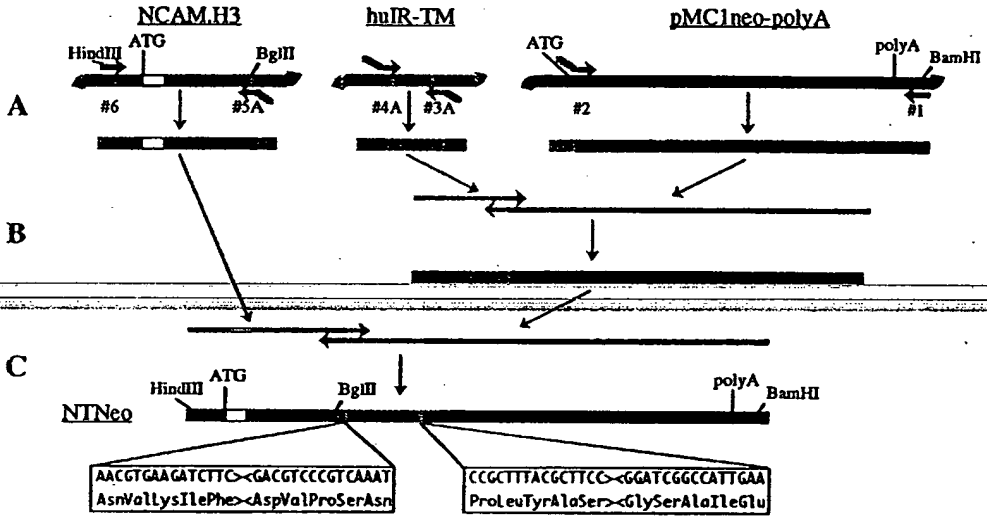


Fig. 1. Construction of the NTNeo insert by recombinant PCR. Thick lines indicate double-stranded template or product. Thin lines represent single-stranded primers or end-annealing complementary template strands. Arrowheads show extension of primers or annealing complementary 3'-ends during PCR. ATG represents the translation start codon; polyA indicates a pair of polyadenylation signals. The three components were first individually amplified by PCR to produce fragments with overlapping sequence (step A). The chimeric gene was produced by annealing and extension of the complementary ends of these fragments, followed by amplification using end-specific primers (steps B and C; primers are not shown for these steps, see Materials and Methods). The huIR-TM (light grey) and *neo* (black) fragments were joined first (step B), with subsequent addition of the 5' NCAM sequence (dark grey; step C), including the signal sequence (white box). The final product, NTNeo is shown with relevant restriction sites and the nucleotide and peptide sequences of in-frame junctions. Linear dimension is not strictly to scale.

Material may be protected by copyright law (Title 17, U.S. Code)

BEST AVAILABLE COPY

the final product. A BglII site in codon 113 of the cDNA is preserved in the NTNeo insert. PCR using primers to huIR produced a 168-bp fragment including the entire 23-codon presumptive transmembrane domain plus 8 and 15 codons of 5'- and 3'-flanking sequence, respectively. The sense primer (#4A) contains a 15-bp NCAM-specific tail and creates the in-frame NCAM/huIR-TM joint; the antisense primer (#3A) has a 15-bp *neo*-specific tail and represents the junction in frame of huIR-TM and *neo*. PCR using primers to pMC1*neo*-polyA produced an 882-bp fragment including bases 735–1597 of the template plasmid. This fragment excludes the start ATG codon but retains all 266 downstream codons of the *neo* gene plus the stop codon and two polyadenylation signals (polyA) from the pMC1*neo*-polyA plasmid template. The sense primer (#2) adds a huIR-TM-specific tail and spans the in-frame union of huIR-TM and *neo* in the final NTNeo insert. The antisense primer (#1) includes a BamHI site from the plasmid template. The following reaction conditions were used: 40 ng/ml template DNA, 500 nM each primer, 250 μM each dNTP, 50 units/ml polymerase, in 1× polymerase buffer (provided with enzyme). $MgCl_2$ concentrations and cycle temperatures varied for each reaction. For #1/#2: 5.5 mM $MgCl_2$; 94°C, 5 min, then 30× (94°C, 1 min; 37°C, 1 min; 72°C, 1 min), then 72°C, 10 min. For #3A/#4A: 1.5 mM $MgCl_2$; 94°C, 5 min, then 30× (94°C, 1 min; 55)C, 1 min; 72°C, 1 min), then 72°C, 10 min. For #5A/#6: 7.5 mM $MgCl_2$; 94°C, 5 min, then 30× (94°C, 1 min; 55°C, 1 min; 72°C, 1 min), then 72°C, 10 min. Products were isolated in low-melting-point agarose gel slices (FMC), melted, diluted into 1 ml of water and used for end-primed recombinant PCR.

Steps B and C in Fig. 1 depict the sequential joining reactions used to create the NTNeo fusion protein. Joining reactions were conducted in three stages to promote specific amplification of joined products, as described below: template cross-priming and extension, low-primer-concentration PCR, and high-primer concentration PCR. In the first stage, cross-priming, two fragments with complementary ends derived from first-round PCR were allowed to anneal slowly and extend in the absence of oligonucleotide primers. The following reaction conditions were used: 12 μl/ml each diluted template, 250 μM each dNTP, 50 units/ml polymerase, 1.5 mM $MgCl_2$, in 1× polymerase buffer. To ensure stringency of primer annealing, cycling was begun with a hot start, by the addition of polymerase to the reaction once all ingredients had reached 94°C. Cycling proceeded as follows: 94°C, 5 min; then 10× (94°C, 1 min; 60°C, 10 min; 72°C, 2 min), then 72°C, 10 min, then 94°C, 5 min. During the last 94°C delay, the second stage of the reaction, low-primer concentration PCR, was begun by the addition of 10 nM each primer and 50 units/ml polymerase in 1× polymerase buffer. This low primer concentration increased the fidelity of amplifying only joined products. Primer pairs used for amplification were complementary to the extreme ends of the joined products: #1/#4A for amplification of the huIR-TM/*neo* product, #1/#6 for the final NTNeo product (primers are not shown for these steps in Fig. 1). PCR proceeded as follows: 20× (94°C, 1 min; 60°C, 2 min; 72°C, 2 min), then 72°C, 10 min, then 94°C, 5 min. During the last 94°C delay, the third stage of the reaction, high-primer concentration PCR, was begun by the addition of 2 μM each primer. This final stage was designed to maximize the quantity of double stranded products. The final amplification proceeded as follows: 5× (94°C, 1 min; 60°C, 2 min; 72°C, 2 min), then 72°C, 10 min.

Plasmid Construction and Sequencing. The final recombinant PCR product, NTNeo, was digested with BamHI and HindIII, gel isolated, and ligated into similarly prepared 3.35-kb plasmid backbone of pSV2*neo* (Fig. 2). The resulting plasmid, pSV2NTNeo,

Fig. 2. Map of pSV2
ligated into the 3.35-
2.4-kb fragment of pS

was transformed
by standard tech
was purified afte
either by double
gradient or by p
ene glycol (15).
frames were con
plasmid with pri
the Sequenase ki
cal).

Cell Culture
cells were maint
fied Eagle's m
supplemented w
and 15% newbo
previously descri
changed every ot
h before transfe
performed by ca
tion (15) using 1
or pSV2neo per
confluence. Cells

Material may be protected by copyright law (Title 17, U.S. Code)

BEST AVAIL

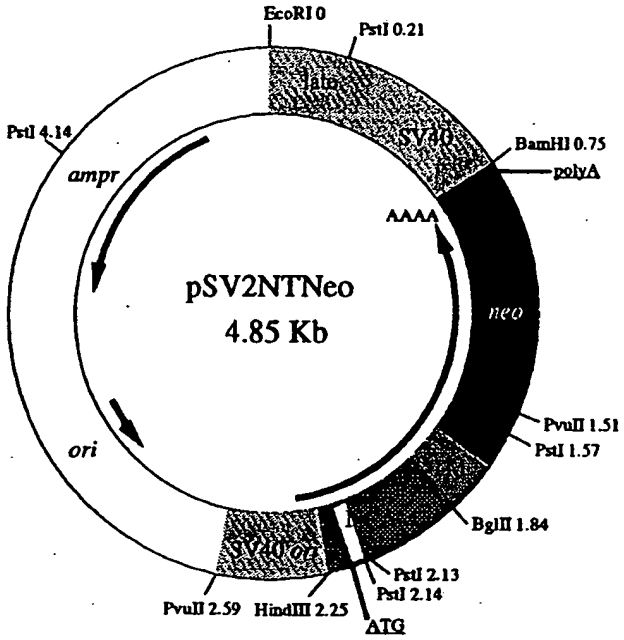


Fig. 2. Map of pSV2NTNeo. The recombinant PCR product NTNeo was digested with HindIII and BamHI and ligated into the 3.35-kb HindIII/BamHI-digested linear backbone of pSV2*neo*. The NTNeo insert replaces the 2.4-kb fragment of pSV2*neo* containing Tn5*neo* and the intron and polyA sites of SV40.

was transformed into bacteria and amplified by standard techniques (15). Plasmid DNA was purified after alkaline lysis of bacteria either by double banding in a CsCl/EtBr gradient or by precipitation with polyethylene glycol (15). Junctions of fused reading frames were confirmed by sequencing of the plasmid with primers #3A and #4A using the Sequenase kit (United States Biochemical).

Cell Culture and Transfection. C2C12 cells were maintained in Dulbecco's modified Eagle's medium (Irvine Scientific) supplemented with 5% fetal bovine serum and 15% newborn calf serum (Hyclone), as previously described (16, 17). Medium was changed every other day, and cells were fed 5 h before transfection. Transfections were performed by calcium phosphate precipitation (15) using 10 μg of either pSV2NTNeo or pSV2neo per 60-mm dish of cells at 20% confluence. Cells were rinsed in phosphate-buffered saline (PBS) and medium was changed after overnight incubation with the precipitate. Two days later cells were passaged to larger dishes, and medium was supplemented with 300 μg/ml G418 (Geneticin, Gibco). Three weeks after the start of drug selection, colonies were ring-cloned by trypsinization and expanded in 300 μg/ml G418.

Cell Extracts and NPT Assays. Assays for NPT activity detected the transfer of radioactive phosphate from ATP to neomycin and were performed as previously described (18), with the modifications provided below. Confluent 10-cm dishes of cells were rinsed with ice-cold phosphate-buffered saline (PBS), and cells were scraped from the dish into 1 ml of cold PBS. After gentle centrifugation, PBS was aspirated, and cell pellets were sonicated in 500 μl of cold homogenization buffer (18) by at least 30 1-sec pulses in an ice-bath sonicator at maximum power (Heat

Material may be protected by copyright law (Title 17, U.S. Code)

System Ultrasonics, model W-225R). Samples were stored on ice as total sonicate. From each sample, 150 μl was removed and spun at 132,000*g* in a Beckman Airfuge for 1 h at 4°C. The top 100 μl of each supernatant was removed and stored on ice as soluble fraction.

NPT assay reactions were carried out in microtiter dishes and contained 67 mM Tris maleate, pH 7.1; 42 mM $MgCl_2$; 400 mM NH_4Cl; 1 μM ATP, 10 Ci/mmol [γ-^{32}P]ATP (Amersham); 500 μM neomycin sulfate (omitted in negative control reactions); plus 5 μl of total sonicate or soluble fraction in a final volume of 55 μl. Reactions proceeded for 1 h at 37°C and were stopped by extraction with a 1:1 mixture of phenol–chloroform. From each aqueous phase, 30 μl was streaked onto P81 cation exchange paper (Whatman) and allowed to air dry. Paper was then washed with continuously flowing tap water for 15 min, and samples were counted for 5 min by Cerenkov emission in a Beckman LS 3801 scintillation counter. As a positive control, pure recombinant NPTII (5Prime-3Prime), which is encoded by the *neo* gene, was added directly to aliquots of the C2C12 control fractions before the start of the NPT assay; the resulting NPTII concentration in the 55 μl reactions was 18 ng/ml.

Protein concentrations in extracts were determined using the Bio-Rad protein assay with bovine serum albumin as a standard.

RESULTS

Construction of Test Plasmid Using Recombinant PCR. To test the possibility of using a membrane-anchored NPT enzyme, we assembled a fusion gene intended to model expression from a hybrid genomic locus created by homologous recombination. Limited by a lack of convenient restriction enzyme sites, we used recombinant PCR to construct a hybrid open reading frame containing elements from three different genes. The PCR construction scheme is shown in Fig. 1 and described in Materials and Methods. Briefly, each of the three portions of the hybrid gene were separately amplified with PCR using primers that resulted in the addition of complementary flanking sequences. These sequences ensured that annealing of the components was specific and in-frame.

The 5′-end portion of the fusion gene comprises sequence from the 5′-untranslated region of a murine neural cell adhesion molecule (NCAM) cDNA clone, designated NCAM.H3 (10), plus the first 115 codons of NCAM coding sequence, which include the translation start codon and *N*-terminal signal sequence. NCAMs comprise a family of proteins that are either secreted or membrane-anchored, via glycophosphatidylinositol linkage or a hydrophobic transmembrane domain. All family members are expressed from a single gene by alternative splicing (19–21). Furthermore, all isoforms of NCAM share the same *N*-terminus, and therefore the same signal sequence. Thus, NCAM serves as a good model for studying problems associated with expression-trap selection for genes encoding secreted, cell surface, and extracellular matrix proteins.

The middle portion of the fusion gene contains the membrane-anchoring domain. For this purpose, we used coding sequence spanning the 23-amino acid transmembrane domain (TM) of human insulin receptor (huIR), plus 23 codons of flanking sequence. This domain of the β-subunit of the receptor is predicted to span the cell membrane, orienting the receptor with the *N*-terminus extracellular and the *C*-terminus cytoplasmic (12).

The 3′-end of the fusion gene construct contains all but the first codon of *neo* from pMC1*neo*-polyA (13), and includes the termination codon, plus 62 bp of additional 3′ untranslated vector sequence containing two

NP
Subst

Fig. 3. Transfection of p
pSV2NTNeo clones equal
produced by transfection o
crude sonicates as describe
in a 1 h reaction. Clones
NTN3) or pSV2*neo* (Neo
omission of the substrate,
specificity of the assay for

polyadenylation signa
genes have been sh
express NPT without
in-frame translation-
sequences (3, 7). In
upstream ATG is th
NCAM. Other than
lies 5′ of the signal
neither the huIR-TM
tion of NCAM incl
contains an in-frame
translation of a fu
without the *N*-termini

The recombinat
inserted into the n
vector pSV2*neo* in pl
to produce pSV2NTN
of the plasmid acros
PCR junctions conf
components of the r
were joined into a sin
frame, as shown in Fi

Material may be protected by copyright law (Title 17, U.S. Code)

BEST AVAILABLE COPY

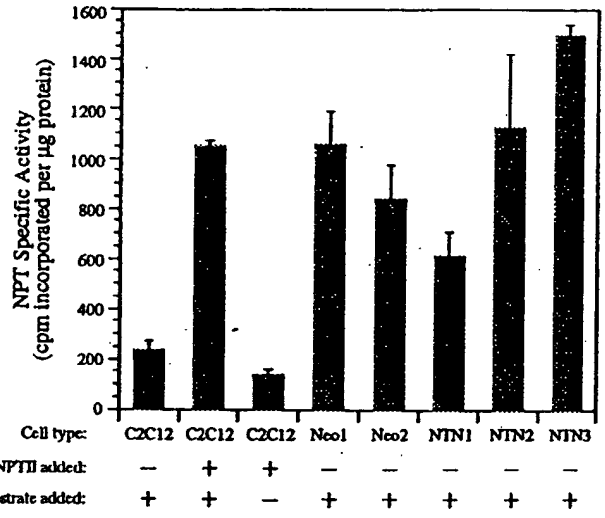


Fig. 3. Transfection of pSV2NTNeo produces NPT activity in cells cloned by drug selection. Activities in pSV2NTNeo clones equal or exceed those for pSV2*neo* clones. Three G418-resistant clones of C2C12 cells, produced by transfection of pSV2NTNeo, are shown compared to controls as described below. NPT was assayed in crude sonicates as described in Materials and Methods, and activity is expressed as cpm incorporated per µg protein in a 1 h reaction. Clones were selected after transfection of C2C12 cells with pSV2NTNeo (NTN1, NTN2, and NTN3) or pSV2*neo* (Neo1, Neo2). Untransfected C2C12 sonicate serves as a negative control. Inclusion and omission of the substrate, neomycin sulfate, from assays of C2C12 sonicate containing pure NPTII demonstrate the specificity of the assay for NPT. Error bars show standard deviation; $N = 3$ for each cell type.

polyadenylation signals. Such ATG-less *neo* genes have been shown to be unable to express NPT without the contribution of an in-frame translation-start site by upstream sequences (3, 7). In NTNeo, the nearest upstream ATG is the start methionine of NCAM. Other than this start codon, which lies 5′ of the signal sequence of NCAM, neither the huIR-TM domain nor the portion of NCAM included in the construct contains an in-frame ATG that would allow translation of a functional *neo* domain without the *N*-terminal signal sequence.

The recombinant PCR product was inserted into the mammalian expression vector pSV2*neo* in place of the Tn*5neo* gene to produce pSV2NTNeo (Fig. 2). Sequencing of the plasmid across the two recombinant PCR junctions confirmed that the three components of the resulting NTNeo insert were joined into a single continuous reading frame, as shown in Fig. 1.

pSV2NTNeo Expresses NPT Activity and Produces Resistance to G418 in Cultured Cells. Transfection of pSV2NTNeo into C2C12 cells reproducibly produced colonies that grew under selection in 300 µg/ml G418. Three representative pSV2NTNeo clones (NTN1, NTN2, NTN3), each produced by a separate transfection experiment, were isolated and used for further study. For comparison, two stable transfectants of pSV2neo (Neo1, Neo2) were also isolated.

To confirm that drug-resistance in pSV2NTNeo clones was due to expression of functional NPT from the vector, assays of NPT activity in pSV2NTNeo clones were compared to assays of pSV2*neo* clones as well as to untransfected C2C12 cells (Fig. 3). All three pSV2NTNeo clones and both pSV2neo clones expressed NPT activity well above background levels in C2C12 cells. Specific activity of NPT in pSV2NTNeo cells was equal to or greater than that seen in

Material may be protected by copyright law (Title 17, U.S. Code)

BEST AVAILABLE COPY

pSV2*neo* cells. Addition of pure NPTII enzyme to C2C12 sonicates indicated that the assay was specifically sensitive to NPT activity in these extracts, as reactions lacking the substrate neomycin sulfate produced only low background signal.

NPT Expressed by pSV2NTNeo is Membrane-Bound. To determine whether the selectable marker expressed from the construct was being directed to the membrane, as hypothesized, we assayed crude and soluble fractions of cell sonicates to test for differences in localization between NPT enzymes expressed from pSV2NTNeo and pSV2neo. Figure 4 shows the relative NPT activities detected in equal volumes of soluble and total extract from each cell type.

Distribution of NPT among the different clones demonstrated clearly that activity derived from pSV2NTNeo is localized in an insoluble compartment, as expected for an integral membrane protein. The total sonicates of all three pSV2NTNeo clones showed higher activity per volume than soluble fractions of the same cell extracts obtained after ultracentrifugation. In contrast, NPT activity in total sonicates of pSV2*neo* cells was markedly reduced relative to the activity assayed in equal volumes of soluble fraction. This result at first appears paradoxical. However, the activity of pure NPTII added to C2C12 cell sonicates gave similar results to those seen for pSV2*neo* cell sonicates. These observations suggest that some insoluble component of the crude sonicates of all cells inhibits assayable NPT activity by approximately 80%, and this inhibition is lost during the 132,000*g* fractionating spin. Therefore, the twofold enrichment of NPT activity seen in total fractions of pSV2NTNeo cells actually indicates a ratio of insoluble to soluble NPT in these extracts of greater than 10 to 1. Combining this evidence with the known functions of the signal sequence and TM domain that were incorporated into NTNeo, we conclude that the fusion protein is localized in the membrane. Moreover, the NPT activity contained in the cytoplasmic

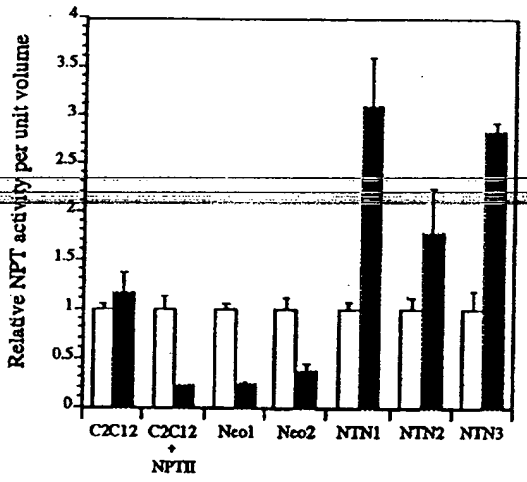


Fig. 4. NPT activity expressed by pSV2NTNeo is insoluble. Assays are of the same cells as shown in Fig. 3. NPT activities in equal volumes of soluble fraction (white bars) or crude sonicate (black bars) are compared for each cell type. In order to compare localization of NPT in different cell types, activity of each fraction for a given clone is normalized to soluble activity in the same clone. NPT activity in pSV2NTNeo cells is clearly lost from sonicates after ultracentrifugation. In contrast, soluble NPT activity in pSV2*neo* cells is markedly increased by the removal of insoluble components. Addition of pure NPTII to control C2C12 fractions confirms the inhibitory nature of crude sonicates. There is no difference in background levels from soluble and crude fractions of untransfected C2C12 cells. Error bars show standard deviation; $N = 3$ for each cell type.

C-terminal domai
of conferring resis
selection.

DISCUSSION

The versatilit
in cytosolic fusio
mented and has e
of *neo*-based expre
targeting (3, 5–8
untranscribed, unt
to select for targ
interest by requir
promoter. A cavea
system on genes co
sequences for vect
membrane. The si
part of the chi
protein expressed
gene. Such a fusion
any membrane anc
exported across th
the drug-degradin
cellular space. Thu
the targeted gen
selecting for rec
intent of the pres
possibility of retai
as a selectable m
including an exoge
main together with
(NTNeo). We hyp
ing chimera would
membrane, with th
in the cytoplasm.
construct, pSV2N
NTNeo fusion pro
enzyme with spe
confers resistance
protein is active
insoluble fraction
tion is presumably
dicted membrane-
component domai
should allow exten
sion-trap-based hc

Material may be protected by copyright law (Title 17, U.S. Code)
BEST AVAILABLE COPY

tracts obtained
contrast, NPT
of pSV2*neo* cells
tive to the activity
oluble fraction.
rs paradoxical.
ure NPTII added
milar results to
nicates. These
some insoluble
nicates of all cells
ty by approxi-
n is lost during
spin. Therefore,
T activity seen
Neo cells actu-
oluble to soluble
ater than 10 to 1.
h the known
ence and TM
ted into NTNeo,
n protein is
Moreover, the
the cytoplasmic

hown in Fig. 3. NPT
mpared for each cell
a given clone is
m sonicates after
by the removal of
tory nature of crude
fected C2C12 cells.

C-terminal domain of the protein is capable of conferring resistance to G418 under drug selection.

DISCUSSION

The versatility of the *neo*-encoded NPT in cytosolic fusion proteins is well documented and has encouraged widespread use of *neo*-based expression-trap vectors for gene targeting (3, 5–8). Such vectors use an untranscribed, untranslatable reporter gene to select for targeting within the gene of interest by requiring transcription from its promoter. A caveat exists for the use of this system on genes coding for *N*-terminal signal sequences for vectorial transport across the membrane. The signal sequence may remain part of the chimeric selectable marker protein expressed from the recombined gene. Such a fusion protein, in the absence of any membrane anchor, would presumably be exported across the plasma membrane and the drug-degrading activity lost to the extracellular space. Thus, the protein encoded by the targeted gene would be useless in selecting for recombination events. The intent of the present study was to test the possibility of retaining the *neo* gene product as a selectable marker within the cell by including an exogenous transmembrane domain together with NPT in a fusion protein (NTNeo). We hypothesized that the resulting chimera would be anchored in the plasma membrane, with the NPT domain suspended in the cytoplasm. Data obtained with our construct, pSV2NTNeo, indicate that the NTNeo fusion protein is indeed a functional enzyme with specific NPT activity that confers resistance to G418. Furthermore, the protein is active in its new location, an insoluble fraction of the cell. This localization is presumably determined by the predicted membrane-targeting function of its component domains. This TM/neo cassette should allow extension of the use of expression-trap-based homologous recombination to a number of genes encoding secreted and membrane-compartmentalized proteins.

A substantial body of work from other researchers demonstrates that expression-trap selection is a reliable, highly efficient approach for detecting homologous recombination into actively expressed target genes (1, 3, 5, 8). Furthermore, although target genes must be expressed in this approach, such selection has been achieved even at loci with very low transcriptional activities (7). The 1.4-kb BglII/BamHI fragment of pSV2NTNeo should serve as a useful cassette in the construction of vectors for the elimination of membrane-targeted proteins via expression-trap homologous recombination. This fragment contains the entire huIR-TM and *neo* domains and could be easily placed at a restriction site within a coding exon of cloned genomic DNA. Unlike insertion of *neo* into the cytoplasmic region of a transmembrane protein (8), which allows expression of more *N*-terminal domains to persist, placement of the TM/*neo* cassette near the 5′-end of the target gene would disrupt expression of functional extracellular domains, as well as the cytoplasmic portions of the targeted protein. Thus, the inclusion of the TM domain in the selectable marker cassette should permit expression-trap disruption of even secreted proteins, while allowing expression of NPT activity within the cytoplasm. The requirement for sufficient 5′ genomic sequence to drive homologous recombination may limit how much of the *N*-terminus of the endogenous protein can be deleted while using expression-trap selection. However, as little as 500 bp of homology in the "short arm" of replacement-type vectors has proven to be enough for homologous recombination (22), and the intron–exon arrangement of many mammalian genes often contains long stretches of gene sequence encoding small amounts of polypeptide. In the case of the NCAM gene, for example, several kilobases of intron sequence exist between regions encoding the

Material may be protected by copyright law (Title 17, U.S. Code)
BEST AVAILABLE COPY

signal sequence and the most *N*-terminal functional domains of the protein (19, 20).

Several other strategies have been used to increase selection efficiencies, even for targeting inactive genes. Addition of negative selectable markers, like herpes simplex thymidine kinase, to the ends of vectors containing a constitutive *neo* gene is the basis for so-called "positive–negative" selection (13, 23, 24). Another approach is the use of a *neo* gene lacking a polyadenylation signal and therefore requiring homologous recombination for proper transcript processing (13). However, although frequencies are enhanced by these approaches, they rarely approach the high rates seen with expression-trap selection on active genes (1, 3, 5, 8). Assuming vectors can be easily constructed using selectable marker cassettes such as the one designed and tested in this study, targeting of active genes by expression trap is likely to provide efficient selection and minimize the requirement for extensive postselection screening of numerous candidate clones.

ACKNOWLEDGMENTS

We thank Dr. Lydia Pan for the NCAM.H3 cDNA, Dr. Richard Roth for the huIR cDNA, and Dr. Lucy Shapiro for NPT assay reagents. We are grateful to Drs. Robert Dubridge, Gary Greenburg, and Gregory Barsh and to members of our laboratory for helpful discussions and critical reading of the manuscript.

We gratefully acknowledge support from a Neonatology and Developmental Biology Training Grant of the National Institutes of Health (HD 07249-11) to W.A.M., and grants from the NIH (HD 18179 and CA 59717), the National Science Foundation (DCB 8417089), and the Muscular Dystrophy Association to H.M.B.

LITERATURE CITED

1. Jasin, M., Elledge, S.J., Davis, R.W., and Berg, P. (1990). *Genes Dev.* 4:157–166.
2. Jasin, M., and Berg, P. (1988). *Genes Dev.* 2:1353–1363.
3. te Riele, H., Maandag, E.R., Clarke, A., Hooper, M., and Berns, A. (1990). *Nature* 348:649–651.
4. Sabatini, D.D., Kreibich, G., Morimoto, T., and Adesnik, M. (1982). *J. Cell Biol.* 92:1–22.
5. Sedivy, J.M., and Sharp, P.A. (1989). *Proc. Natl. Acad. Sci. U.S.A.* 86:227–231.
6. Schwartz, F., Maeda, N., Smithies, O., Hickey, R., Edelmann, W., Skoultchi, A., and Kucherlapati, R. (1991). *Proc. Natl. Acad. Sci. U.S.A.* 88:10416–10420.
7. Jeannotte, L., Ruiz, J.C., and Robertson, E.J. (1991). *Mol. Cell. Biol.* 11:5578–5585.
8. Arbonés, M.L., Austin, H.A., Capon, D.J., and Greenburg, G. (1994). *Nature Genet.* 6:90–97.
9. Fire, A., Harrison, S.W., and Dixon, D. (1990). *Gene* 93:189–198.
10. Pan, L.C., Margolskee, R.F., and Blau, H.M. (1992). *Somat. Cell Mol. Genet.* 18:163–177.
11. Santoni, M.J., Barthels, D., Vopper, G., Boned, A., Goridis, C., and Wille, W. (1989). *Mol. Cell. Biol.* 10:1969–1981.
12. Ebina, Y., Ellis, L., Jarnagin, K., Edery, M., Graf, L., Clauser, E., Ou, J.-h., Marsiarz, F., Kan, Y.W., Goldfine, I.D., Roth, R.A., and Rutter, W.J. (1985). *Cell* 40:747–758.
13. Thomas, K.R., and Capecchi, M.R. (1987). *Cell* 51:503–512.
14. Southern, P.J., and Berg, P. (1982). *J. Mol. Appl. Genet.* 1:327–341.
15. Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., and Struhl, K. (1993). *Current Protocols in Molecular Biology* (John Wiley & Sons, New York).
16. Yaffe, D., and Saxel, O. (1977). *Nature* 270:725–727.
17. Blau, H.M., Chiu, C.P., and Webster, C. (1983). *Cell* 32:1171–1180.
18. Ramesh, N., and Osborne, W.R.A. (1991). *Anal. Biochem.* 193:316–318.
19. Hirsch, M.-R., Gaugler, L., Deagostini-Bazin, H., Bally-Cuif, L., and Goridis, C. (1990). *Mol. Cell. Biol.* 10:1959–1968.
20. Cunningham, B.A., Hemperly, J.J., Murray, B.A., Prediger, E.A., Brackenbury, R., and Edelman, G.M. (1987). *Science* 236:799–806.
21. Barbas, J.A., Chaix, J.-C., Steinmetz, M., and Goridis, C. (1988). *EMBO J.* 7:625–632.
22. Hasty, P., Rivera-Pérez, J., and Bradley, A. (1991). *Mol. Cell. Biol.* 11:5586–5591.
23. Tybulewicz, V.L., Crawford, C.E., Jackson, P.K., Bronson, R.T., and Mulligan, R.C. (1991). *Cell* 65:1153–1163.
24. Mansour, S.L., Thomas, K.R., and Capecchi, M.R. (1988). *Nature* 336:348–352.

Material may be protected by copyright law (Title 17, U.S. Code)

BEST AVAILABLE COPY